CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/70, 1/68, C12N 15/00, A61K 39/29 — A1

(11) International Publication Number: WO 99/51781

(43) International Publication Date: 14 October 1999 (14.10.99)

(21) International Application Number: PCT/US99/07404

(22) International Filing Date: 2 April 1999 (02.04.99)

(30) Priority Data:

60/080,509	2 April 1998 (02.04.98)	US
60/090,356	23 June 1998 (23.06.98)	US

(71) Applicant *(for all designated States except US)*: VIROPHARMA INCORPORATED [US/US]; 405 Eagleview Boulevard, Exton, PA 19341 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: COLLETT, Marc, S. [US/US]; 113 Juniper Court, Collegeville, PA 19426 (US).

(74) Agents: RIGAUT, Kathleen, D. et al.; Dann, Dorfman, Herrell and Skillman, Suite 720, 1601 Market Street, Philadelphia, PA 19103 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: HEPATITIS C VIRUS NS5B COMPOSITIONS AND METHODS OF USE THEREOF

(57) Abstract

HCV NS5B nucleic acids and proteins are disclosed herein. Also provided are methods and kits utilizing the compositions of the invention.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Hepatitis C Virus NS5B Compositions and Methods of Use Thereof

Field of the Invention

The present invention relates to hepatitis C virus (HCV), and more specifically, to the HCV NS5B gene and the encoded protein. In particular, the invention relates to novel compositions comprising HCV NS5B sequences, functional HCV NS5B sequences, functionally improved compositions of HCV NS5B sequences, and to the use of such NS5B sequences in research, diagnostic, therapeutic and pharmaceutical applications.

Cross Reference to Related Applications

This application claims priority under 35 U.S.C. §119(e) to US Provisional Application 60/080,509 filed April 2, 1998 and US Provisional Application 60/090,356 filed June 23, 1998, the disclosures of each being incorporated by reference as though set forth herein in full.

Background of the Invention

Hepatitis C virus (HCV) is a major cause of hepatitis globally. The World Health Organization estimates that 150 million people worldwide are presently infected with the virus. Most infections become persistent and about 60% of cases develop into chronic liver disease. Chronic HCV infection can lead to development of cirrhosis, hepatocellular carcinoma and liver failure.

For treatment of hepatitis due to HCV, interferon

alpha (IFN) is currently approved in the U.S. IFN treatment is associated with improved serum enzyme response in 20-40% of patients. The remainder are nonresponsive to IFN treatment. For responders, a sustained improvement of aminotransferase levels is seen in only 10-20% of patients; the majority of patients relapse upon cessation of IFN treatment. While IFN represents the first treatment of chronic hepatitis C, its effectiveness is variable, its cure rate is low, and associated adverse effects are considerable.

Vaccines under development for HCV generally consist of recombinant versions of the putative viral structural proteins (C, E1, E2), or genes encoding these. It is believed that virus neutralizing antibodies do exist, can be elicited, and may be able to inhibit or prevent HCV infection. However, to date, no vaccine has been demonstrated safe and effective for HCV. Indeed, given the inherent genetic diversity of HCV, with virus isolates exhibiting immunologically distinct envelope proteins that are not neutralized by pre-existing antibodies, vaccine development will be a formidable task.

HCV can not yet be efficiently propagated *in vitro*. This deficiency has compromised a clear understanding of several aspects of HCV replication. While numerous HCV isolates have been molecularly cloned from infected patients and have been sequenced, much remains to be learned regarding the required sequence features and elements which facilitate efficient replication. From study of these sequences in a variety of surrogate gene expression systems, knowledge of the molecular biology

of HCV has expanded greatly in recent years. These advancements, which have been facilitated by existing knowledge and understanding of the molecular biology of the related pestiviruses and flaviviruses, have yielded important insights into the virus-specified components essential for virus replication that may be suitable targets for antiviral drug discovery strategies.

As mentioned above, HCV is a genetically heterogeneous virus. It exists in nature, and in infected individuals, as a "quasispecies", which means that within the virus population from any particular source, the viral genetic material is present as a collection of closely related, but non-identical sequences. This genetic heterogeneity of HCV sequences may be a consequence of natural errors introduced during the normal replication process of HCV, of selection pressures placed on the virus by the host, or others factors. It is believed that the quasispecies nature of HCV may contribute to the virus's ability to avoid elimination by the host's defense systems and establish persistent infections.

The quasispecies nature of HCV has made it difficult to define the sequence of "functional" genes and genomes, particularly in the absence of a controlled system for the efficient replication of the virus. However, functionality of segments of HCV sequences has been studied in a variety of surrogate expression systems. Several gene products have been shown to be functional in *in vitro* assays in that they exhibit enzymological activities that are predicted to exist based on the presence of particular amino acid sequence

motifs in these gene products. For example, the NS3 protein possesses serine proteinase, nucleoside triphosphatase and helicase activities and the NS5B protein possesses RNA-dependent RNA polymerase (RdRp) activity.

The recent demonstration of an infectious molecular clone of HCV [Kolykhalov *et al.*, Science 277:570-574 (1997); Yanagi *et al.*, Proc. Natl. Acad. Sci. USA 94:8738-8743 (1997)] allows definition of a complete functional HCV genome and consequently of a sequence for a set of functional HCV gene products from this particular strain of HCV obtained from one particular source. It is noteworthy that Kolykhalov *et al.* indicate that the difficulties in obtaining a functional genome included: 1) the highly variable quasispecies nature of HCV in infected patients; 2) the required *in vitro* amplification (cDNA synthesis and PCR amplification) of the sequences due to the small quantities of viral RNA in clinical samples; and 3) the necessity of molecular cloning (in *E. coli*). Each of these aspects allows for the possibility and presentation of defective (poorly functional or nonfunctional) HCV sequences or for errors in transcribing or amplifying HCV sequences. Kolykhalov *et al.* indicated that RNAs produced from 34 full genome molecular clones failed to exhibit infectivity in chimpanzees. Nucleotide sequence analyses showed numerous sequence changes among the 6 full length clones that were completely sequenced. Only after a "consensus sequence" clone was generated, was positive infectivity demonstrable. Yanagi *et al.* also indicated the high

degree of sequence variability among individual molecular clones obtained from a single source material and the non-infectious nature of RNA transcripts from full-length molecular clones. Yanagi *et al.* concluded that a large proportion of HCV genomes are defective. In fact, recent work indicates a bias toward the preferential selection of defective HCV genomes during molecular cloning procedures [Forns, *et al.*, Proc Natl Acad Sci USA 94:13909-13914 (1997)]. Of 25 random clones derived from a control plasmid by Forns *et al.*, only 8% were functional in polyprotein synthesis.

Thus, it is clear from these examples that mere knowledge of a sequence derived from HCV is insufficient to conclude that such a sequence encodes useful viral genes or functional viral gene products.

Regarding the NS5B gene, while many partial and complete gene sequences have been entered into various databases, only a very few of these sequences have been demonstrated to be functional based on *in vitro* assays of RdRp activity.

The present invention provides novel and functional, as well as functionally improved, hepatitis C virus NS5B sequences for use in research, diagnostic, therapeutic and pharmaceutical applications, and for use in assays for the identification of efficacious antiviral agents.

SUMMARY OF THE INVENTION

The present invention provides novel HCV NS5B nucleotide sequences from which recombinant HCV proteins

having demonstrable RdRp activity may be derived. Sequence modifications that result in improved functional activity are also provided. Such functional recombinant HCV NS5B proteins have utility in research, diagnostic, therapeutic and pharmaceutical applications. In particular, the recombinant HCV NS5B proteins of the invention have utility in antiviral drug discovery strategies.

In a preferred embodiment of the invention, an isolated nucleic acid molecule is provided that comprises a DNA sequence identified as SEQ ID NO: 1, which is present in clone 4 of the present invention. An exemplary HCV NS5B protein has the amino acid sequence identified as SEQ ID NO: 2 encoded by clone 4. Additional nucleic acid molecules that represent nucleic acid sequences related to those of SEQ ID NO: 1, including but not limited to those identified by SEQ ID NO: 3, 4, 5, 6 and 7 encoding, respectively, the amino acid sequences identified by SEQ ID NO: 8, 9, 10, 11 and 12 of the HCV NS5B are also contemplated to be within the scope of the present invention. It is further contemplated that conservative sequence or residue substitutions of these sequences are also within the scope of the invention.

In an additional embodiment of the invention, functionally improved hepatitis C polymerase nucleic acid and amino acid sequences are provided. Exemplary sequences include those nucleic acid sequences that encode amino acid sequences related to those identified

by SEQ ID NO: 2 in which particular codons have been substituted to encode particular amino acid changes in the sequence. Examples of useful changes include but are not limited to changes at amino acid positions number 75, number 177 and number 543 of the mature NS5B protein.

In another embodiment of the invention, functionally improved hepatitis C polymerase nucleic acid and amino acid sequences are provided in which particular codons are substituted to encode particular amino acids not normally found in known sequences of NS5B. For example, changes including but not limited to residue changes at amino acid positions number 1 and number 2 of the mature NS5B protein are useful according to this invention.

As described above, nucleic acids encoding variant proteins or polypeptides are contemplated to be within the scope of the present invention. Sequence ID NO: 13 provides an example of such a variant sequence. Such variants may or may not possess HCV polymerase activity. These variants may possess one or more changes each of which may include one or more additions, deletions, or substitutions of amino acid residues. Preferably, the changes will not affect, or substantially affect, the structure or useful properties of the polypeptide. Thus, HCV NS5B variants may suitably possess functional NS5B activity such as those described above, or they may be poorly functional or inactive, yet contain substantially the secondary and tertiary structure of the native polypeptide. Such NS5B molecules may be used to advantage to identify agents that specifically bind

to or otherwise affect the HCV NS5B activity. HCV NS5B variants can be either naturally occurring (i.e., purified or isolated from a natural source) or synthetic (i.e., generated by biological expression of DNA that has been subjected to site-directed mutagenesis or produced by chemical synthetic techniques well known in the art).

The nucleic acid molecules of the invention may be cloned and expressed in vectors. Such vectors may be in the form of, for example, a plasmid, a replication competent or defective virus or phage vector or a replicon provided typically with an origin of replication, optionally a promoter for the expression of the polynucleotide and optionally a regulator of the promoter. The vector may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid or a neomycin resistance gene for a mammalian vector. The vector may be used *in vitro*, for example for the production of RNA or protein. The vector may be further used to transform, transfect, infect or transduce a host cell or an organism. The present invention further contemplates the use of host cells and organisms harboring or expressing the HCV nucleic acid sequences or polypeptides of the invention for the identification of agents that affect the activity of the HCV NS5B protein.

In yet another embodiment of the invention, methods are provided for the identification of agents that affect the HCV NS5B polymerase sequences. Such methods include high throughput screening procedures that allow assessment of large numbers of agents. The agents

identified by use of the HCV NS5B nucleic acids and polypeptides of the invention, variants thereof, or methods of the invention, may be either antagonistic or agonistic in their affect on the NS5B sequences. These agents may include molecules of any number of classes including but not limited to small molecules, polymers, peptides, polypeptides, immunoglobulins or fragments thereof, oligonucleotides, antisense molecules, , peptide-nucleic acid conjugates, ribozymes, polynucleotides and the like. It is specifically contemplated that both antagonistic and agonistic molecules identified by practice of the invention have broad and multiple utilities. Such utilities for antagonists of HCV NS5B activity include, but are not limited to, uses for the inhibition of HCV replication in humans, in other living hosts and in *in vitro* systems such as cell, tissue and organ cultures. Agonists of HCV NS5B activity identified by practice of the invention will also have multiple utilities, both in living hosts and in *in vitro* systems. For example, such agents will be useful in the development of animal models of HCV infection, replication or disease and for the propagation of HCV in a living host or in cell, tissue or organ culture systems.

According to another aspect of the invention, kits are provided to facilitate the use of the compositions and methods disclosed herein. Exemplary kits would include HCV NS5B nucleic acids and polypeptides of the invention, variants thereof, alone or in association with suitable vectors. Also included would be protocols for use of the compositions of the invention for the

particular application and the necessary reagents to carry out the application. The reagents of a kit may vary depending on the intended application. Such reagents may include, but are not limited to buffers, solvents, media and solutions, substrates and cofactors, vectors and host cells, and detection or reporter reagents. Other accessories may also be included such as vials, vessels and reaction chambers.

The following definitions are provided to aid in understanding the subject matter regarded as the invention.

As used herein, "hepatitis C virus" or "HCV" shall mean any representative of a diverse group of related viruses classified within the hepacivirus genus of the Flaviviridae family.

"Nucleic acid" or a "nucleic acid molecule" as used herein refers to any DNA or RNA molecule, either single or double stranded and, if single stranded, the molecule of its complementary sequence in either linear or circular form. In discussing nucleic acid molecules, a sequence or structure of a particular nucleic acid molecule may be described herein according to the normal convention of providing the sequence in the 5' to 3' direction. With reference to nucleic acids of the invention, the term "isolated nucleic acid" is sometimes used. This term, when applied to DNA, refers to a DNA molecule that is separated from sequences with which it is immediately contiguous in the naturally occurring genome of the organism in which it originated. For example, an "isolated nucleic acid" may comprise a DNA molecule inserted into a vector, such as a plasmid or

virus vector, or integrated into the genomic DNA of a prokaryotic or eukaryotic cell or host organism.

When applied to RNA, the term "isolated nucleic acid" refers primarily to an RNA molecule encoded by an isolated DNA molecule as defined above. Alternatively, the term may refer to an RNA molecule that has been sufficiently separated from other nucleic acids with which it would be associated in its natural state (i.e., in cells or tissues). An isolated nucleic acid (either DNA or RNA) may further represent a molecule produced directly by biological or synthetic means and separated from other components present during its production.

"Natural allelic variants", "mutants" and "derivatives" of particular sequences of nucleic acids refer to nucleic acid sequences that are closely related to a particular sequence but which may possess, either naturally or by design, changes in sequence or structure. By closely related, it is meant that at least about 75%, but often, more than 90%, of the nucleotides of the sequence match over the defined length of the nucleic acid sequence referred to using a specific SEQ ID NO. Changes or differences in nucleotide sequence between closely related nucleic acid sequences may represent nucleotide changes in the sequence that arise during the course of normal replication or duplication in nature of the particular nucleic acid sequence. Other changes may be specifically designed and introduced into the sequence for specific purposes, such as to change an amino acid codon or sequence in a regulatory region of the nucleic acid. Such specific changes may be made *in vitro* using

a variety of mutagenesis techniques or produced in a host organism placed under particular selection conditions that induce or select for the changes. Such sequence variants generated specifically may be referred to as "mutants" or "derivatives" of the original sequence.

The terms "percent similarity", "percent identity" and "percent homology" when referring to a particular sequence are used as set forth in the University of Wisconsin GCG software program.

The term "NS5B" refers to a portion of the HCV genome located near the 3' end of the viral genome that specifies the region encoding a protein, termed the "NS5B protein", "NS5B polypeptide", "NS5B polymerase" or combinations of these terms which are used interchangeably herein. NS5B in its natural state, functions as an RNA-dependent RNA polymerase (RdRp). The nucleic acid region encoding the NS5B protein may also be referred to as the "NS5B gene". Thus, the term "NS5B" may refer to either a nucleic acid encoding the NS5B polypeptide, to an NS5B gene or to an NS5B polypeptide, or to any portions thereof, depending on the context in which the term is used. NS5B may further refer to natural allelic variants, mutants and derivatives of either NS5B nucleic acid sequences or NS5B polypeptides. The NS5B nucleic acid, NS5B gene or NS5B protein referred to may be either functional or non-functional.

The present invention also includes active portions, fragments, derivatives and functional or non-functional mimetics of HCV NS5B polypeptides or

proteins of the invention. An "active portion" of HCV NS5B polypeptide means a peptide that is less than the full length HCV NS5B polypeptide, but which retains measurable biological activity.

A "fragment" or "portion" of the HCV NS5B polypeptide means a stretch of amino acid residues of at least about five to seven contiguous amino acids, often at least about seven to nine contiguous amino acids, typically at least about nine to thirteen contiguous amino acids and, most preferably, at least about twenty to thirty or more contiguous amino acids. Fragments of the HCV NS5B polypeptide sequence, antigenic determinants, viral antigens or epitopes are useful for eliciting immune responses to a portion of the HCV NS5B amino acid sequence.

A "derivative" of the HCV NS5B polypeptide or a fragment thereof means a polypeptide modified by varying the amino acid sequence of the protein, e.g. by manipulation of the nucleic acid encoding the protein or by altering the protein itself. Such derivatives of the natural amino acid sequence may involve insertion, addition, deletion or substitution of one or more amino acids, and may or may not alter the essential activity of original the HCV NS5B polypeptide. As mentioned above, the HCV NS5B polypeptide or protein of the invention includes any analogue, fragment, derivative or mutant which is derived from a HCV NS5B polypeptide and which retains at least one property or other characteristic of the HCV NS5B polypeptide. Different "variants" of the HCV NS5B polypeptide exist in nature. These variants may be alleles characterized

by differences in the nucleotide sequences of the gene coding for the protein, or may involve different RNA processing or post-translational modifications. The skilled person can produce variants having single or multiple amino acid substitutions, deletions, additions or replacements. These variants may include inter alia: (a) variants in which one or more amino acids residues are substituted with conservative or non-conservative amino acids, (b) variants in which one or more amino acids are added to the HCV NS5B polypeptide, (c) variants in which one or more amino acids include a substituent group, and (d) variants in which the HCV NS5B polypeptide is fused with another peptide or polypeptide such as a fusion partner, a protein tag or other chemical moiety, that may confer useful properties to the HCV NS5B polypeptide, such as, for example, an epitope for an antibody, a polyhistidine sequence, a biotin moiety and the like. Other HCV NS5B polypeptides of the invention include variants in which amino acid residues from one species are substituted for the corresponding residue in another species, either at the conserved or non-conserved positions. In another embodiment, amino acid residues at non-conserved positions are substituted with conservative or non-conservative residues. The techniques for obtaining these variants, including genetic (suppressions, deletions, mutations, etc.), chemical, and enzymatic techniques are known to the person having ordinary skill in the art.

To the extent such allelic variations, analogues, fragments, derivatives, mutants, and modifications,

including alternative nucleic acid processing forms and alternative post-translational modification forms result in derivatives of the HCV NS5B polypeptide that retain any of the biological properties of the HCV NS5B polypeptide, they are included within the scope of this invention.

The term "functional" as used herein implies that the nucleic or amino acid sequence is functional for the recited assay or purpose.

The phrase "consisting essentially of" when referring to a particular nucleotide or amino acid means a sequence having the properties of a given SEQ ID No:. For example, when used in reference to an amino acid sequence, the phrase includes the sequence per se and molecular modifications that would not affect the basic and novel characteristics of the sequence.

A "replicon" is any genetic element, for example, a plasmid, cosmid, bacmid, phage or virus, that is capable of replication largely under its own control. A replicon may be either RNA or DNA and may be single or double stranded.

A "vector" is a replicon, such as a plasmid, cosmid, bacmid, phage or virus, to which another genetic sequence or element (either DNA or RNA) may be attached so as to bring about the replication of the attached sequence or element.

An "expression operon" refers to a nucleic acid segment that may possess transcriptional and translational control sequences, such as promoters, enhancers, translational start signals (e.g., ATG or AUG codons), polyadenylation signals, terminators, and the

like, and which facilitate the expression of a polypeptide coding sequence in a host cell or organism.

The term "oligonucleotide," as used herein refers to primers and probes of the present invention, and is defined as a nucleic acid molecule comprised of two or more ribo- or deoxyribonucleotides, preferably more than three. The exact size of the oligonucleotide will depend on various factors and on the particular application and use of the oligonucleotide.

The term "probe" as used herein refers to an oligonucleotide, polynucleotide or nucleic acid, either RNA or DNA, whether occurring naturally as in a purified restriction enzyme digest or produced synthetically, which is capable of annealing with or specifically hybridizing to a nucleic acid with sequences complementary to the probe. A probe may be either single-stranded or double-stranded. The exact length of the probe will depend upon many factors, including temperature, source of probe and use of the method. For example, for diagnostic applications, depending on the complexity of the target sequence, the oligonucleotide probe typically contains 15-25 or more nucleotides, although it may contain fewer nucleotides. The probes herein are selected to be "substantially" complementary to different strands of a particular target nucleic acid sequence. This means that the probes must be sufficiently complementary so as to be able to "specifically hybridize" or anneal with their respective target strands under a set of pre-determined conditions. Therefore, the probe sequence need not reflect the exact complementary sequence of the target. For example, a

non-complementary nucleotide fragment may be attached to the 5' or 3' end of the probe, with the remainder of the probe sequence being complementary to the target strand. Alternatively, non-complementary bases or longer sequences can be interspersed into the probe, provided that the probe sequence has sufficient complementarity with the sequence of the target nucleic acid to anneal therewith specfically.

The term "specifically hybridize" refers to the association between two single-stranded nucleic acid molecules of sufficiently complementary sequence to permit such hybridization under pre-determined conditions generally used in the art (sometimes termed "substantially complementary"). In particular, the term refers to hybridization of an oligonucleotide with a substantially complementary sequence contained within a single-stranded DNA or RNA molecule of the invention, to the substantial exclusion of hybridization of the oligonucleotide with single-stranded nucleic acids of non-complementary sequence.

The term "primer" as used herein refers to an oligonucleotide, either RNA or DNA, either single-stranded or double-stranded, either derived from a biological system, generated by restriction enzyme digestion, or produced synthetically which, when placed in the proper environment, is able to functionally act as an initiator of template-dependent nucleic acid synthesis. When presented with an appropriate nucleic acid template, suitable nucleoside triphosphate precursors of nucleic acids, a polymerase enzyme, suitable cofactors and conditions such as a suitable

temperature and pH, the primer may be extended at its 3' terminus by the addition of nucleotides by the action of a polymerase or similar activity to yield an primer extension product. The primer may vary in length depending on the particular conditions and requirement of the application. For example, in diagnostic applications, the oligonucleotide primer is typically 15-25 or more nucleotides in length. The primer must be of sufficient complementarity to the desired template to prime the synthesis of the desired extension product, that is, to be able anneal with the desired template strand in a manner sufficient to provide the 3' hydroxyl moiety of the primer in appropriate juxtaposition for use in the initiation of synthesis by a polymerase or similar enzyme. It is not required that the primer sequence represent an exact complement of the desired template. For example, a non-complementary nucleotide sequence may be attached to the 5' end of an otherwise complementary primer. Alternatively, non-complementary bases may be interspersed within the oligonucleotide primer sequence, provided that the primer sequence has sufficient complementarity with the sequence of the desired template strand to functionally provide a template-primer complex for the synthesis of the extension product. Amino acid residues described herein are preferred to be in the "L" isomeric form. However, residues in the "D" isomeric form may be substituted for any L-amino acid residue, provided the desired properties of the polypeptide are retained.

All amino-acid residue sequences represented herein conform to the conventional left-to-right amino-terminus to carboxy-terminus orientation.

The term "isolated protein" or "isolated and purified protein" is sometimes used herein. This term refers primarily to a protein produced by expression of an isolated nucleic acid molecule of the invention. Alternatively, this term may refer to a protein that has been sufficiently separated from other proteins with which it would naturally be associated, so as to exist in "substantially pure" form. "Isolated" is not meant to exclude artificial or synthetic mixtures with other compounds or materials, or the presence of impurities that do not interfere with the fundamental activity, and that may be present, for example, due to incomplete purification, addition of stabilizers, or compounding into, for example, immunogenic preparations or pharmaceutically acceptable preparations.

The term "substantially pure" refers to a preparation comprising at least 50-60% by weight of a given material (e.g., nucleic acid, oligonucleotide, protein, etc.). More preferably, the preparation comprises at least 75% by weight, and most preferably 90-95% by weight of the given compound. Purity is measured by methods appropriate for the given compound (e.g. chromatographic methods, agarose or polyacrylamide gel electrophoresis, HPLC analysis, and the like).

"Mature protein" or "mature polypeptide" shall mean a polypeptide possessing the sequence of the polypeptide after any processing events that normally occur to the polypeptide during the course of its genesis, such as

protoelytic processing from a polyprotein precursor. In designating the sequence or boundaries of a mature protein, the first amino of the mature protein sequence is designated as amino acid residue 1. In the case of the mature NS5B protein, its normal biogenesis entails its proteolytic cleavage from a precursor polyprotein. Thus, the first amino acid residue of the mature NS5B protein represents a serine residue for all known HCV sequences. As used herein, any amino acid residues associated with a mature protein not naturally found associated with that protein that precedes amino acid 1 are designated amino acid -1, -2, -3 and so on. For recombinant expression systems, a methionine initiator codon is often utilized for purposes of efficient translation. Thus, in the case of NS5B, a methionine codon may be place immediately proceeding the serine codon. This methionine residue in the resulting polypeptide, as used herein, would be positioned at -1 relative to the mature NS5B protein sequence.

The term "tag," "tag sequence" or "protein tag" refers to a chemical moiety, either a nucleotide, oligonucleotide, polynucleotide or an amino acid, peptide or protein or other chemical, that when added to another sequence, provides additional utility or confers useful properties, particularly in the detection or isolation, to that sequence. Thus, for example, a homopolymer nucleic acid sequence or a nucleic acid sequence complementary to a capture oligonucleotide may be added to a primer or probe sequence to facilitate the subsequent isolation of an extension product or hybridized product. In the case of protein tags,

histidine residues (e.g., 4 to 8 consecutive histidine residues) may be added to either the amino- or carboxy-terminus of a protein to facilitate protein isolation by chelating metal chromatography. Alternatively, amino acid sequences, peptides, proteins or fusion partners representing epitopes or binding determinants reactive with specific antibody molecules or other molecules (e.g., flag epitope, c-myc epitope, transmembrane epitope of the influenza A virus hemaglutinin protein, protein A, cellulose binding domain, calmodulin binding protein, maltose binding protein, chitin binding domain, glutathione S-transferase, and the like) may be added to proteins to facilitate protein isolation by procedures such as affinity or immunoaffinity chromatography. Chemical tag moieties include such molecules as biotin, which may be added to either nucleic acids or proteins and facilitates isolation or detection by interaction with avidin reagents, and the like. Numerous other tag moieties are known to, and can be envisioned by, the trained artisan, and are contemplated to be within the scope of this definition.

As used herein, the terms "reporter," "reporter system", "reporter gene," or "reporter gene product" shall mean an operative genetic system in which a nucleic acid comprises a gene that encodes a product that when expressed produces a reporter signal that is a readily measurable, e.g., by biological assay, immunoassay, radioimmunoassay, or by colorimetric, fluorogenic, chemiluminescent or other methods. The nucleic acid may be either RNA or DNA, linear or

circular, single or double stranded, antisense or sense polarity, and is operatively linked to the necessary control elements for the expression of the reporter gene product. The required control elements will vary according to the nature of the reporter system and whether the reporter gene is in the form of DNA or RNA, but may include, but not be limited to, such elements as promoters, enhancers, translational control sequences, poly A addition signals, transcriptional termination signals and the like.

The terms "transform", "transfect", "transduce", shall refer to any method or means by which a nucleic acid is introduced into a cell or host organism and may be used interchangeably to convey the same meaning. Such methods include, but are not limited to, transfection, electroporation, microinjection, PEG-fusion and the like.

The introduced nucleic acid may or may not be integrated (covalently linked) into nucleic acid of the recipient cell or organism. In bacterial, yeast, plant and mammalian cells, for example, the introduced nucleic acid may be maintained as an episomal element or independent replicon such as a plasmid. Alternatively, the introduced nucleic acid may become integrated into the nucleic acid of the recipient cell or organism and be stably maintained in that cell or organism and further passed on or inherited to progeny cells or organisms of the recipient cell or organism. In other manners, the introduced nucleic acid may exist in the recipient cell or host organism only transiently.

A "clone" or "clonal cell population" is a

population of cells derived from a single cell or common ancestor by mitosis.

A "cell line" is a clone of a primary cell or cell population that is capable of stable growth *in vitro* for many generations.

An "immune response" signifies any reaction produced by an antigen, such as a viral antigen, in a host having a functioning immune system. Immune responses may be either humoral in nature, that is, involve production of immunoglobulins or antibodies, or cellular in nature, involving various types of B and T lymphocytes, dendritic cells, macrophages, antigen presenting cells and the like, or both. Immune responses may also involve the production or elaboration of various effector molecules such as cytokines, lymphokines and the like. Immune responses may be measured both in *in vitro* and in various cellular or animal systems. Such immune responses may be important in protecting the host from disease and may be used prophylactically and therapeutically.

A "viral antigen" shall be any peptide, polypeptide or protein sequence, segment or epitope that is derived from a virus that has the potential to cause a functioning immune system of a host to react to said viral antigen.

An "antibody" or "antibody molecule" is any immunoglobulin, including antibodies and fragments thereof, that binds to a specific antigen. The term includes polyclonal, monoclonal, chimeric, and bispecific antibodies. As used herein, antibody or antibody molecule contemplates both an intact

immunoglobulin molecule and an immunologically active portion of an immunloglobulin molecule such as those portions known in the art as Fab, Fab', F(ab')2 and F(v).

As used herein, the term "living host" shall mean any non-human autonomous being.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows an alignment of the amino acid sequences of several NS5B proteins of the invention. The approximate positions of conserved sequence motifs present in RdRp enzymes identified by Koonin [J Gen Virol. (1991) 72:2197-206] (Roman numerals) and Poch *et al.* [EMBO J. (1989) 8:3867-74] (letters) are indicated above the clone 4 sequence.

Figure 2 shows the purification of the HCV NS5B protein. Sf9 cells infected with a recombinant baculovirus expressing the HCV NS5B gene were harvested and lysed. Cleared lysate was subjected to a series of protein purification steps. Samples for each step of the purification procedure were electrophoresed on an SDS-containing polyacrylamide gel, which was then silver-stained. Lane 1: cleared cell lysate; lane 2, flow-through material from DEAE column; lane 3, material from heparin column; lane 4, material from Cibracron blue column. M = molecular mass standards; numbers in kilodaltons.

Figure 3 is an alignment of the clone 4 NS5B sequence of the invention with the genotype 1a consensus

NS5B sequence derived from an infectious clone (GenBank Accession Number AF009606) and a genotype 1a NS5B sequence (PCT WO 97/12033) previously shown to have RdRp activity. The dashed lines indicate amino acid residue identity.

Figure 4 shows an alignment of the clone 4 NS5B sequence of the invention with a genotype 1b "consensus" amino acid sequence derived from sequences of HCV NS5B sequences obtained from the GenBank database and with genotype 1b NS5B sequences previously shown to be functional (GenBank Accession Numbers M58335 and Z97730). The dashed lines indicate amino acid residue identity.

Figure 5 shows a protein sequence alignment of the HCV genotype 1b NS5B gene reported by Yanagi *et al.* [Virology (1998) 244:161-172; GenBank Accession Number AF054247] and clone 4 of the present invention. The dashed lines indicate amino acid residue identity.

Figure 6 shows results of a high throughput assay for HCV NS5B RdRp activity.

Figure 7 shows the sequence of SEQ ID NO: 1.

Figure 8 shows the sequence of SEQ ID NO: 3.

Figure 9 shows the sequence of SEQ ID NO: 4.

Figure 10 shows the sequence of SEQ ID NO: 5.

Figure 11 shows the sequence of SEQ ID NO: 6.

Figure 12 shows the sequence of SEQ ID NO: 7.

Figure 13 shows the amino acid sequence of clone 20(V-1M), SEQ ID NO: 13.

DETAILED DESCRIPTION OF THE INVENTION

HCV is the major causative agent of transfusion-associated and sporadic non-A, non-B hepatitis. A high number of HCV-infected patients develop chronic hepatitis that eventually leads to cirrhosis and often progresses to hepatocellular carcinoma. There is an urgent need for new effective treatments of this disease.

HCV is an enveloped virus with a (+) stranded linear RNA genome of approximately 9.4 kilobases. This nucleic acid encodes a large polyprotein that is processed by viral and cellular proteases into at least 9 different viral polypeptides.

In one of its aspects, the present invention provides HCV genetic material derived from an HCV-infected patient. The invention further provides HCV recombinant NS5B proteins expressed from these sequences and sequences that have RdRp activity.

Expression of recombinant HCV NS5B gene sequences may be carried out in a variety of systems including but not limited to bacterial, yeast, mammalian, insect and plant cell systems, as well as in organisms such as infected, transfected, transduced or transgenic insects, animals or plants. In one embodiment of the invention,

recombinant baculoviruses were constructed to express HCV NS5B gene sequences in insect cells following infection in culture.

RNA extracted from the serum of a patient infected with HCV was subjected to a reverse transcriptase-nested polymerase chain reaction procedure using primer oligonucleotides designed to amplify the NS5B gene sequences. Primers for the nested PCR reaction allowed direct cloning of the NS5B gene into a baculovirus expression transfer vector, providing an initiator methonine codon immediately preceeding the first amino acid of the authentic NS5B coding sequence. The NS5B gene from several transfer vector clones was sequenced and used to generate recombinant baculoviruses. Infection of Sf9 cells with these viruses, followed by Western immunoblotting with HCV NS5B sequence-specific antiserum, showed clear production of the 68 kilodalton NS5B protein.

In another aspect of the invention, NS5B sequences that are expressed, purified and evaluated for RdRp activity may possess varying levels of RdRp activity, from highly active to little or no activity, and may thus represent functional, poorly functional or non-functional sequences, respectively. Such sequences would also be assessed for activity in accordance with the present invention.

In a further aspect of the invention, HCV NS5B proteins may be modified by particular changes in nucleotide and amino acid sequence that result in RdRp enzymes with altered functionality. Such changes may be subtle and represent conservative substitutions such as

in the case of nucleotide sequences, changes in the codon sequence that do or do not alter the encoded amino acid, or for amino acid sequences, changes that result in conservative residue substitutions, additions or deletions.

In yet a further aspect of the invention, NS5B nucleic acids and polypeptides derived from sequences of the invention have utility in numerous methods, assays and kits involving research, diagnostic, therapeutic and pharmaceutical applications, and in the development of antiviral strategies for the prevention and treatment of HCV disease.

Based on the discoveries described herein, it is demonstrated that simple knowledge of, or mere inspection of, a nucleotide or amino acid sequence of an NS5B gene or protein is insufficient to enable one to deduce or predict that the particular NS5B sequence represents a sequence that has enzymatic activity or functional utility.

Preparation of HCV NS5B Nucleic Acid Molecules and HCV NS5B Proteins and Uses Thereof in Assay Methods and Kits

A. <u>Nucleic Acid Molecules</u>

Nucleic acid molecules encoding the HCV NS5B proteins of the invention may be prepared by two general methods: (1) They may be synthesized from appropriate chemical starting materials, or (2) they may be isolated from biological sources. Both methods utilize protocols well known in the art.

The availability of nucleotide sequence

information, such as that provided herein for HCV NS5B sequences, enables preparation of an isolated nucleic acid molecule of the invention by oligonucleotide synthesis. Synthetic oligonucleotides may be prepared by the phosphoramadite method employed in the Applied Biosystems 38A DNA Synthesizer or similar devices. The resultant construct may be purified according to methods known in the art, such as high performance liquid chromatography (HPLC). Long, double-stranded polynucleotides, such as a DNA molecule of the present invention, must be synthesized in stages due to the size limitations inherent in current oligonucleotide synthetic methods. Thus, for example, a 3 kilobase double-stranded molecule may be synthesized as several smaller segments of appropriate complementarity. Complementary segments thus produced may be ligated such that each segment possesses appropriate cohesive termini for attachment of an adjacent segment. Adjacent segments may be ligated by annealing cohesive termini in the presence of DNA ligase to construct an entire 3 kilobase double-stranded molecule. A synthetic DNA molecule so constructed may then be cloned and amplified in an appropriate vector.

Nucleic acid sequences encoding HCV NS5B proteins may be isolated from appropriate biological sources using methods known in the art. For example, RNA isolated from the serum of an HCV infected patient may be used as a suitable starting material for the generation of cDNA molecules encoding HCV NS5B proteins.

In accordance with the present invention, nucleic acids having the appropriate level of sequence homology

with the protein coding region of the DNA molecules of the present invention may be identified by using hybridization and washing conditions of appropriate stringency. For example, hybridizations may be performed, using a hybridization solution comprising, for example, 5X SSC, 5X Denhardt's reagent, 1.0% SDS, 100 μg/ml denatured, fragmented salmon sperm DNA, 0.05% sodium pyrophosphate and up to 50% formamide. Hybridization is carried out at 37-42°C for at least six hours. Following hybridization, filters are washed as follows: (1) 5 minutes at room temperature in 2X SSC and 1% SDS; (2) 15 minutes at room temperature in 2X SSC and 0.1% SDS; (3) 30 minutes-1 hour at 37°C in 1X SSC and 1% SDS; (4) 2 hours at 42-65°C in 1X SSC and 1% SDS, changing the solution every 30 minutes.

One common formula for calculating the stringency conditions required to achieve hybridization between nucleic acid molecules of a specified sequence homology is as follows (Sambrook et al., 1989):

$T_m = 81.5°C + 16.6Log\ [Na+] + 0.41(\%\ G+C) - 0.63\ (\%\ formamide) - 600/\#bp\ in\ duplex$

As an illustration of the above formula, using [Na+] = [0.368] and 50% formamide, with GC content of 42% and an average probe size of 200 bases, the T_m is 57°C. The T_m of a DNA duplex decreases by 1-1.5°C with every 1% decrease in homology. Thus, targets with greater than about 75% sequence identity would be observed using a hybridization temperature of 42°C. Such a sequence would be considered substantially homologous to the sequences of the present invention.

Nucleic acids of the invention may be maintained as DNA in any convenient cloning vector. In one embodiment, clones are maintained in plasmid cloning/expression vectors, such as pBluescript plasmids (Stratagene, La Jolla, CA) or recombinant baculovirus transfer vectors such as pFastBac vectors (Gibco-BRL, Gaithersburg, MD) that are propagated in suitable *E. coli* host cells.

The nucleic acids of the invention may also be used as starting materials for the generation of sequence variants of the nucleic acids of the invention using any number of synthetic and molecular biologic procedures well known in the art including but not limited to site-directed mutagenesis techniques. Particular mutations may give rise to HCV NS5B proteins with altered characteristics such as increased enzymatic activity.

HCV NS5B protein-encoding nucleic acid molecules of the invention include cDNA, genomic DNA, RNA, and fragments thereof, which may be single- or double-stranded in nature. Thus, this invention provides oligonucleotides (sense or antisense strands of DNA or RNA) having sequences capable of hybridizing with at least one sequence of a nucleic acid molecule of the present invention, such as selected segments of the cDNA having substantially the sequence of any of the sequences identified in the present invention. Such oligonucleotides are further useful as probes and primers for detecting or isolating additional HCV NS5B encoding nucleic acids.

B. Proteins

HCV NS5B proteins of the present invention may be prepared in a variety of ways, according to any number of known methods. The protein may be purified from appropriate sources, e.g., cultured cells, tissues or organs, by a variety of techniques that may include partitioning and precipitation procedures, affinity purification methods, conventional chromatography procedures, high performance chromatography techniques and the like.

The availability of nucleic acids molecules encoding HCV NS5B protein enables production of the protein using *in vitro* expression methods known in the art. For example, a cDNA or gene may be cloned into an appropriate *in vitro* transcription vector, such as pSP64 or pSP65 for *in vitro* RNA synthesis, followed by cell-free translation of the RNA in a suitable cell-free translation system, such as extracts of wheat germ, rabbit reticulocytes or HeLa cells. *In vitro* transcription and translation systems are commercially available (e.g., Promega Biotech, Madison, WI; Gibco-BRL, Gaithersburg, MD).

Alternatively, according to a preferred embodiment of the invention, larger quantities of HCV NS5B protein may be produced by expression in suitable prokaryotic or eukaryotic systems such as bacterial, fungal, mammalian or plant systems. For example, part or all of a DNA molecule, such as a cDNA may be inserted into a plasmid vector adapted for expression in a bacterial cell, such as *E. coli*, or into a baculovirus vector for expression in an insect cell. Such vectors comprise the regulatory

elements necessary for expression of the DNA in the host cell (e.g., *E. coli* or insect cell), positioned in such a manner as to permit expression of the DNA in the host cell. Such regulatory elements required for expression may include promoter sequences, transcriptional initiation and termination sequences, enhancer sequences, translational control sequences and the like.

The HCV NS5B proteins or derivatives thereof produced by gene expression in a recombinant prokaryotic or eukaryotic system may be purified according to methods known in the art. In one embodiment, a commercially available expression/secretion system can be used, whereby the recombinant protein is expressed and thereafter secreted from the host cell, to be easily purified from the surrounding medium. If expression/secretion vectors are not used, an alternative approach involves purifying the recombinant protein from extracts of expressing cells, tissues or organs by standard protein purification techniques or by affinity separation techniques, such as by immunological interaction with antibodies that bind specifically to the recombinant protein or by nickel columns for isolation of recombinant proteins tagged with 5-8 histidine residues at their N-terminus or C-terminus. Such methods are commonly used by skilled practitioners.

The HCV NS5B proteins of the invention, prepared by the aforementioned methods, may be analyzed according to standard procedures. For example, such proteins may be subjected to electrophoretic analyses and to amino acid sequence analyses, as well as to crystallographic analyses for structure determination according to known

methods. Such analyses provide useful information regarding the functionality of the NS5B protein and on means to affect that functionality, such as in the design of molecules that may inhibit the function of the NS5B protein.

C. Assay Methods and Kits

The HCV NS5B sequences of the invention may be used in a variety of ways having utility in research, diagnostic, therapeutic and pharmaceutical applications. Representative methods of use for the compositions of the invention are described below.

In one aspect, the nucleic acid sequences of the invention, and sequences complementary to these, may be used as probes or primers for the detection, labeling, identification or isolation of related nucleic acids in biological or synthetic preparations. For example, nucleic acid sequences of the invention may be used as hybridization probes to detect the presence of HCV in samples. Such hybridization probes may further be used to isolate the nucleic acids to which they are hybridized by techniques well known in the art. Additionally, the nucleic acid sequences of the invention may be used as primers for the detection or isolation of HCV or related nucleic acids using techniques such as reverse transcriptase-polymerase chain reaction (RT-PCR). Appropriate primers pairs may be further used in nested PCR applications. Such primers, primer pairs and probes may represent any

portion of the NS5B sequences of the invention. The actual sequence of the NS5B gene used will vary according to the specific application. Moreover, additional sequences may be added to the HCV primer or probe sequence, such as homopolymer tails (tags), sequences that represent useful restriction enzyme recognition sites, sequences encoding particular amino acid residues, initiation or termination codons or other sequences that may be useful for the particular application at hand. Typically, oligonucleotides of from 10 to 80 nucleotides in length that are either the same as or complementary to the sequences of the invention are useful as hybridization probes or as primers in RT-PCR applications. Alternatively, the entire NS5B sequence may be employed as a capture hybridization probe, for example.

Several examples of useful oligonucleotide primers and probes of the invention comprise sequences set forth in Table 1, sequences complementary to these and potions of these sequences.

Table 1

Oligonucleotides useful as probes and primers

5'-TCAATGTCCTACACATGGAC-3'	SEQ ID NO: 14
5'-CTACACATGG-3'	SEQ ID NO: 15
5'-CTCTGATTACACCATGCGCTGCGGAGGAGAGCAAGCTGCCC-3'	SEQ ID NO: 16
5'-AATGCGCTGAGCAACTCTTTGCTGCGCC-3'	SEQ ID NO: 17
5'-CCATAACATGGTCTATGCCACAACATCCCGCAGCGCAAGCCAGCGGC-3'	SEQ ID NO: 18
5'-GAAGAAGGTAACTTTTGACAGG-3'	SEQ ID NO: 19
5'-CAAGTCCTGGATGACCACTACCG-3'	SEQ ID NO: 20
5'-GACGTGCTCAAGGA-3'	SEQ ID NO: 21
5'-ATGAAGGCGAAGGCGTCC-3'	SEQ ID NO: 22
5'-GGAAGAAGCCTGTAAG-3'	SEQ ID NO: 23
5'-GAACCTATCCAGCAAGGCCGTTAA-3'	SEQ ID NO: 24

5'-ACCAGAGAAAGGAGGCCGC-3'	SEQ ID NO: 25
5'-ACCCAGACTTGGGG-3'	SEQ ID NO: 26
5'-TCTCCACCCTTCCTCAGGCT-3'	SEQ ID NO: 27
5'-CGAGTTCCTGGTGAATGCC-3'	SEQ ID NO: 28
5'-TGCCCTATGGGCTTCGCATATGAC-3'	SEQ ID NO: 29
5'-TTTCGACTCAACGGTCACCGAGAAT-3'	SEQ ID NO: 30
5'-GTTGAGGAGTCAATT-3'	SEQ ID NO: 31
5'-TTGGCCCCCGAAGCCAGACA-3'	SEQ ID NO: 32
5'-AAGGTCGCTTACAGAGC-3'	SEQ ID NO: 33
5'-ATCGGGGGTCCCCTGAC-3'	SEQ ID NO: 34
5'-TAACTCAAAAGGGCAGAG-3'	SEQ ID NO: 35
5'-ATGTTACTTGAAGGCCTCT-3'	SEQ ID NO: 36
5'-GATGCTTGTGTGCGGAGACGACCTC-3'	SEQ ID NO: 37
5'-GGTCGCGCACGATGCATCTGGCAAAAGGGTA-3'	SEQ ID NO: 38
5'-CACCACCCCTCTTGCGCGG-3'	SEQ ID NO: 39
5'-CTCCATCCTTCTAGCTCAGGAGCAA-3'	SEQ ID NO: 40
5'-AGTTACTGTCCCAGGGGGGG-3'	SEQ ID NO: 41
5'-TCCGGCTGCGTCCCAGT-3'	SEQ ID NO: 42
5'-CCGACCCCGCTGGTTCATGTGGTGCC-3'	SEQ ID NO: 43
5'-CTACCTGCTCCCGAACCGA-3'	SEQ ID NO: 44
5'-CTACCTGCTCCCCAACCGA-3'	SEQ ID NO: 45

Additionally, the nucleic acid sequences of the invention may be used as primers for the generation of variants or mutants of the sequences of the invention using a variety of methodologies known in the art, including site-directed mutagenesis procedures.

In another aspect, the nucleic acid sequences of the invention may be used in the construction or generation of, or incorporated into, infectious viruses, vectors or replicons. Provision or substitution of the functionally superior NS5B sequences of the invention for poorly functional or non-functional counterparts will serve to improve the infectious and replicative

characteristics of the resulting virus or replicating unit. For example, Rice *et al.* (PCT WO 98/39031) provide an infectious nucleic acid of HCV. The NS5B polymerase encoded by this infectious clone is shown in the present invention to be substantially inferior in its RdRp activity to those of the present invention. Therefore, substitution of the NS5B encoding gene of Rice *et al.* with those of the present invention would be expected to result in improved virus replication due to the improved characteristics of the NS5B gene and encoded protein of the present invention. Such substitutions may be carried out by standard genetic engineering procedures well known in the art. The resulting infectious nucleic acid would have considerable advantages over current infectious clones of HCV, including, but not limited to, improved or higher level viral RNA synthesis, improved levels of infectious virus production and improved virus replication in living hosts and in *in vitro* systems. Use in other systems in which the function of the HCV NS5B polymerase is important, such as in complementing or trans-complementing systems, replicon systems, defective viruses, defective interfering particles and the like, would also benefit from the use of the nucleic acid sequences of the invention.

In another embodiment, the nucleic acid sequences of the invention may be used in methods to elicit immune responses to the HCV NS5B protein. For example, the NS5B-encoding nucleic acid sequences of the invention operationally linked to an expression operon may introduced directly into cells, particularly into

antigen presenting cells such as dendritic cells, of a living host or human possessing a functioning immune system. Introduction of the sequences may utilize transfection, transformation, or transduction methods, or involve the physical uptake of particles coated with the NS5B nucleic acid sequences, such as plasmid-coated gold particles. Once inside cells, the NS5B sequences are expressed, processed and presented to the host's immune system. Such methods are useful in the elicitation of humoral and cellular immune responses in a living host and in humans and in vaccines for HCV.

The nucleic acid sequences of the invention may be further used in the generation of cell lines or cellular systems that express the HCV NS5B protein. Such cell lines in which a functional NS5B protein is expressed from the NS5B genes of the invention will have utility in methods for assaying materials for antagonistic or agonistic activity toward HCV. For example, assays may be established whereby intact cells expressing an HCV NS5B protein of the invention are contacted with agents or materials suspected of affecting the intracellular activity of the HCV NS5B protein, and the affect of such agents on the HCV NS5B activity is measured. The affect of such agents on the HCV NS5B activity may be measured in any number of ways. For example, RNA synthesis that is directly or indirectly dependent on the HCV NS5B polymerase activity may be quantified. In one manner, the amount of a radiolabeled precursor of RNA (e.g., ^{3}H-uridine) that is incorporated into trichloroacetic acid-precipitable RNA that is dependent on NS5B activity may be measured.

Alternatively, such cell systems may utilize a reporter system in which the production of the reporter signal is dependent on RNA synthesis by the HCV NS5B polymerase. In one embodiment, a RNA substrate of the HCV NS5B polymerase is provided that is the antisense strand of an mRNA, the sense strand (mRNA) of which is effectively translated to produce a polypeptide capable of being detected or of producing a detectable signal (the reporter). For example, an RNA molecule is provided comprising the sequence complementary to the coding sequence of luciferase (antisense strand). The activity of the HCV NS5B polymerase on this RNA results in the production of the sense strand of the luciferase gene, which is then translated by the cellular translational system to produce luciferase protein. The luciferase protein then may be detected by antibodies to the luciferase protein or by measurement of luciferase enzymatic activity in intact cells or in cellular extracts using a luminometer or other similar device. Numerous other reporters may serve equally well in this application including but not limited to, β-galactosidase, alkaline phosphatase, fluorescent green protein and the like. Furthermore, the cell systems that may be used in this method of the invention may be of bacterial, fungal, insect, avian, mammalian or plant origin.

Further, the nucleic acid sequences of the invention may be used in assays to identify agents or materials capable of interacting with or affecting the HCV NS5B nucleic acid sequences. For example, assays

may be established in which nucleic acid sequences of the invention are provided and then contacted with agents or materials suspected of interacting with such sequences. Agents identified in such interaction assays would then have potential diagnostic utility and uses in the detection of HCV in, for example, biological samples. Such agents would also have potential utility in applications involving the prevention or treatment of HCV disease in an affected living host, including humans, and for the inhibition or enhancement of HCV replication or propagation in living hosts and in *in vitro* systems such as cell, tissue and organ cultures. Additional applications may be envisioned once the nature of the particular agent is clear.

The HCV NS5B protein compositions of the invention also have broad utility. In diagnostic applications, for example, the NS5B proteins, or peptides thereof, may be used in assays for the detection of immune responses to the same. For example, protein sequences or peptides of the invention may be used in assays in which said sequences are immobilized on a matrix and used to capture antibodies directed to said sequences. Additionally, protein sequences or peptides of the invention may be used to detect or measure cell-mediated immune responses to the protein, such as in immune cell proliferation assays.

The HCV NS5B protein compositions of the invention also have potential utility in the elicitation of immune

responses, such as in vaccines. For example, provision of the NS5B proteins of the invention, or peptides thereof, to a living organism with a functioning immune system will cause such organism to mount an immune response to the NS5B sequences. The NS5B sequences may be presented to the living organism in any number of assays well known to those trained in the art and include, but are not limited to, providing free protein or peptides, formulated protein or peptides, adjuvanted protein or peptides, protein or peptides in the context of intact or disrupted cells in which NS5B sequences are present and other such manners. Immune responses so elicited may be either humoral or cellular in nature, or both. Such immune responses may be important in protecting living hosts from HCV disease and may also provide or serve as a source of useful immunological reagents such as antibodies, that may further have therapeutic or diagnostic utility. The NS5B proteins of the invention, or peptide portions thereof, may further be used to select or purify such antibodies to NS5B. For example, NS5B protein may be immobilized and used to bind antibodies specific to the NS5B protein, and thus enrich for such antibodies. Furthermore, the NS5B proteins and peptides of the invention may be used to produce monoclonal antibodies to the NS5B protein using standard techniques known in the art. Antibodies to the NS5B protein, whether polyclonal or monoclonal, may be further evaluated for their ability to affect the

enzymatic activity of the NS5B RdRp activity.

Several examples of portions of the NS5B protein sequences useful in the above applications comprise sequences or portions thereof, provided in Table 2.

Table 2

Peptides Useful in Methods of the Invention

SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASQRQK KVTFDRLQVLDDHYRDVLKEMKAKASTVKA	SEQ ID NO: 46
SVEEACKLTPPHSARSKFGYGAKDVRNLSSKAVNHIHSVWKDLLEDT	SEQ ID NO: 47
ETPIDTTIMAKNEVFCVQPEKGGRKPARLIVYPDLGVRVC	SEQ ID NO: 48
SSYGFQYSPGQRVEFLVNAWKSKKCPMGFAYDTRCFDSTVTENDIRVEESIY QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQSCGYRRCRASGVLTTSCG	SEQ ID NO: 49
CTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLELITSC	SEQ ID NO: 50
SVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNS	SEQ ID NO: 51
CLRKLGVPPLRVWRHRARSVRAKLLSQGG	SEQ ID NO: 52
FVAGYSGGDIYHSLSRARPRWFMWC	SEQ ID NO: 53

The protein compositions of the invention have utility in assays for the detection and identification of agents capable of interacting with or affecting the HCV NS5B protein. Assays may be established in which HCV NS5B polypeptide sequences of the invention are provided and then contacted with agents or materials suspected of interacting with such sequences. For example, upon provision of the HCV NS5B protein of the invention, or fragment or portion thereof, contacted

agents may be assessed for their ability to bind specifically to the protein. Such binding agents would then have potential diagnostic utility and uses in the detection of HCV in, for example, biological samples. Such binding agents may further affect the functional activity of the HCV protein, such as either inhibiting or enhancing the HCV NS5B function. Agents that inhibit the function of the NS5B protein would have potential utility in applications involving the prevention or treatment of HCV disease in an affected living host or for the inhibition of HCV replication or propagation in living hosts, including humans, in *in vitro* systems such as cell, tissue and organ cultures or in biological materials. Agents that enhance the function of the NS5B protein would have potential utility in applications involving the replication, propagation or production of HCV in living hosts, such as in animal models of HCV replication, and in *in vitro* systems such as cell, tissue and organ cultures.

In another embodiment, methods of assay are provided in which the HCV NS5B polymerase activity furnished by an enzymatically active NS5B protein of the invention is measured directly. Agents placed in contact with said enzymatically active NS5B polymerase may be assessed for their ability to specifically affect this enzymatic activity. Such enzymatically active polymerase may be provided in an extract or lysate of a cell in which the polypeptide was produced, in an *in*

vitro cell-free expression system or in an enriched or purified form.

There are numerous means by which the enzymatic activity of the HCV NS5B protein provided in an extract, cell-free system or enriched form may be assessed, and these are well known in the art. NS5B-dependent RNA synthesis typically requires certain reaction components including minimally a buffered medium, a divalent cation, precursors of RNA (nucleoside triphosphates, NTPs), an RNA template and a primer for RNA synthesis on that template. Additional components may include monovalent cations, reducing agents, stabilizers, cofactors and inhibitors of activities unrelated to the NS5B RdRp activity such as inhibitors of RNase, phosphatase, kinase, phosphotransferase and similar activities.

Measurement of NS5B-dependent RNA synthesis may be assessed in numerous manners. In one example, the incorporation of a precursor of RNA into a polymer of RNA is measured, such as the incorporation of a radiolabeled NTP into trichloroacetic acid-precipitable RNA, which may then be quantified by scintillation spectrometry or phosphorimaging technologies. Such precursors may alternatively be tagged with other moieties to allow their ready detection such as with biotin for detection with avidin reagents including various avidin conjugates such alkaline phosphatase and the like or with fluorescently-labeled NTPs for

detection using fluorescent technologies such as fluorescence polarization.

NS5B-dependent RNA synthesis may also be assessed by measuring the extension of a pre-labeled or tagged primer of RNA synthesis such as a radiolabeled or biotin-tagged oligonucleotide that is used by the polymerase to initiate RNA synthesis on a template RNA molecule. Extension of the primer may be assessed by quantifying the addition of nucleoside triphosphates to the primer, by determining the length of the primer product, or by other methods known in the art.

Alternatively, the product of NS5B RdRp activity may be detected and quantified by capture of the product RNA using hybridization techniques. For example, an oligonucleotide complementary to the product of the NS5B RdRp reaction may be introduced during or after the reaction and hybridized to the product. The extent of hybridization of the added oligonucleotide may be used as a measure of the amount of product RNA present in the mixture and may be assessed by various means known in the art.

Other means of detection of the products of NS5B RdRp activity are readily known to, or can be envisioned by, the skilled artisan and are fully contemplated here.

Assays involving the nucleic acid and polypeptide compositions of the invention may be formatted in any number of configurations. Particularly useful for evaluating large numbers of agents and materials are high throughput screening formats. Traditionally such assays were typically formatted in 96 well plates. However, 384, 864 and 1536 well plates may be used in

such high throughput assay systems. These systems are often automated using robotics technologies to allow manipulation and processing of large numbers of samples.

The agents or materials that may be evaluated in the various assay methods of the invention for potential antagonistic or agonistic affects include but are not limited to small molecules, polymers, peptides, polypeptides, proteins, immunoglobulins or fragments thereof, oligonucleotides, antisense molecules, peptide-nucleic acid conjugates, ribozymes, polynucleotides and the like. The potential utility of agents or materials identified using the compositions and assay methods of the invention will be broad and will include uses for the detection and isolation of HCV nucleic acids and polypeptides, for the detection or diagnosis of HCV, for the prevention and treatment of HCV disease in an affected living host, including humans, and for the inhibition or enhancement of HCV replication or propagation in living hosts and in *in vitro* systems such as cell, tissue and organ cultures, as well as for other uses the may be envisioned once the nature of the agent is clear.

Another feature of the invention includes kits to facilitate the use of the compositions and methods disclosed herein. Exemplary kits would include HCV NS5B nucleic acids and polypeptides of the invention, and/or variants thereof, alone or in suitable vectors. Also included would be protocols for use of the compositions of the invention for the particular application and the necessary reagents to carry out the application. Such reagents may include, but not be limited to, buffers,

solvents, media and solutions, substrates and cofactors, vectors and host cells, and detection or reporter reagents. Accessory items may include vials, vessels, reaction chambers and instruction sheets.

The following examples are provided to describe the invention in further detail. These examples, which set forth the preferred mode presently contemplated for carrying out the invention, are intended to illustrate and not to limit the invention.

EXAMPLE 1

Cloning and Expression of HCV NS5B Genes

RNA isolated from the serum of an HCV-infected patient was used to amplify the HCV NS5B gene using an RT-nested PCR protocol. Using the following first round primers: 5'-TGA GGA TGT CGT CTG CTG CTC AAT GTC C-3' and 5'-GGG ATG GCC TAT TGG CCT GGA GT-3', the RT reaction was performed at 50°C for 50 minutes followed by a 1 minute 94°C denaturation, followed by 40 cycles of first round PCR, 15 cycles of 94°C for 30 seconds, 50°C for 30 seconds 68°C for 2 minutes and 25 cycles of 94°C for 30 seconds, 50°C for 30 seconds, 68°C for 2 minutes.

A portion of first round PCR reaction mixture was then used in the second round PCR reaction in which the following nested primers were incorporated: 5'-AAC AGA TCT GAA TTC TTA TAA ATA TGT CAA TGT CCT ACA CAT GGA C-3' and 5'-TGC TCT AGA GCG GCC GCT CAT CAT CGG TTG GGG AGC AGG TAG-3', which included EcoRI and NotI restriction

sites, respectively (underlined) for subsequent cloning purposes. The second round PCR involved an initial denaturation at 94°C for 1 minute, followed by 10 cycles of 94°C for 45 seconds, 50°C for 30 seconds, 68°C for 2 minutes and 20 cycles of 94°C for 30 seconds, 50°C for 30 seconds, 68°C for 2 minutes.

The resultant PCR product was purified, digested with EcoRI and NotI and ligated to pFastBac plasmid (Gibco-BRL) previously digested with EcoRI and NotI. The ligation mixture was transformed into DH5 *E. coli* cells. Plasmids from bacterial colonies containing the HCV NS5B gene were used to generate a recombinant baculovirus by the transposition method of the Bac-to-Bac system (Gibco-BRL). After transformation of the HCV NS5B gene-containing pFastBac plasmid DNA into *E. coli* DH10Bac cells, several colonies containing bacmid DNA were transfected into Sf9 insect cells according to the protocol supplied by the manufacturer. NS5B protein expression in recombinant baculovirus-infected cells was verified by Western immunoblot analysis with antiserum specific to the HCV NS5B sequences.

Six independently isolated clones (clones 4, 14, 21, 11, 16 and 20) of the HCV NS5B gene from a single RNA preparation derived from one patient were sequenced. The nucleotide sequences of these clones are provided by SEQ ID NO: 1, 3, 4, 5, 6 and 7, respectively. All six sequences were derived from a genotype 1b virus and are closely related, but not identical, to one another. Each of the sequences set forth above is unique and novel sequence, not presently listed in the GenBank

database. The deduced amino acid sequences of nucleotide sequences identified by SEQ ID NO: 1, 3, 4, 5, 6 and 7 are provided by SEQ ID NO: 2, 8, 9, 10, 11 and 12, respectively. These NS5B amino acid sequences are aligned relative to clone 4 (SEQ ID NO: 2) in Figure 1. As revealed in this alignment, the sequence of the NS5B protein of clones 14, 21, 11, 16 and 20 differs from that of clone 4 by 2, 5, 3, 4 and 3 amino acids, respectively.

EXAMPLE 2

Purification of HCV NS5B Proteins

HCV NS5B proteins may be obtained in purified form from recombinant protein-expressing cell systems by any number of procedures known to the trained artisan, several of which are exemplified in Current Protocols in Molecular Biology, Frederick M. Ausubel *et al.* eds., John Wiley & Sons, 1995 which is incorporated by reference herein.

In the case of the present example, baculovirus clone 4-infected insect cells were disrupted in lysis buffer (50% glycerol, 20 mM Tris-HCl, pH 7.5, 10 mM dithiothreitol (DTT), 0.5 M NaCl, 1 mM EDTA, 2% Triton X-100) supplemented with a Complete Protease Inhibitor tablet (Boehringer Mannheim). $MgCl_2$ was then added to a final concentration of 10 mM, followed by 10 units DNase I (RQ-1, Promega). After 30 minutes on ice, the lysate

was clarified by centrifugation at 35,000 rpm for 35 minutes at 4°C.

The clarified lysate was diluted to a final NaCl concentration of 0.3 M with elution buffer (20% glycerol, 20 mM Tris-HCl, pH 7.5, 10 mM DTT, 1 mM EDTA, 0.5% Triton X-100), and incubated with DEAE Sepharose equilibrated in elution buffer containing 0.3 M NaCl at 4°C. The mixture was then poured into a column, and flow-through material was collected. Flow-through material was diluted to a final NaCl concentration of 0.2 M with elution buffer and loaded onto a heparin Sepharose column equilibrated in elution buffer containing 0.2 M NaCl. Bound proteins were eluted with a linear gradient of NaCl (200 mM to 1 M NaCl in elution buffer) while fractions were collected.

NS5B-containing fractions were pooled and loaded onto a Cibacron Blue column equilibrated in elution buffer containing 0.4 M NaCl. Bound proteins were eluted with a linear gradient of NaCl (400 mM to 4M NaCl in elution buffer) while fractions were collected. NS5B-containing fractions were pooled and dialyzed against 50% glycerol, 10 mM Tris-HCl, pH 7.2, 50 mM NaCl, 1 mM EDTA, 0.01% Triton X-100 and stored at -20°C. Typical purification results are shown in Figure 2.

EXAMPLE 3

Enzymatic Activity of Purified HCV NS5B Proteins

There are numerous methodologies for the measurement of RNA-dependent RNA polymerase activity

that are well known to one of ordinary skill in the art.

One approach for measuring the HCV NS5B RdRp activity uses a purified recombinant NS5B protein in an *in vitro* RdRp assay. For example, Behrens *et al.* [EMBO J. 15:12-22 (1996)] describe the baculovirus expression, purification and enzymatic activity of the HCV NS5B RdRp derived from the BK strain of HCV. In another example, PCT WO 97/12033 [PCT/US96/15571] discloses the bacterial expression, purification and enzymatic activity of the HCV NS5B RdRp derived from an HCV sample obtained from the Centers for Disease Control and several truncated and modified versions of this sequence. Another example is Lohmann *et al.* [J Virol 71:8416-8428 (1997)], in which the NS5B gene derived from a chronically infected patient was expressed with recombinant baculoviruses in insect cells, purified and enzymatically evaluated. Yet another example is Yuan *et al.* [Biochem Biophys Res Comm 232:231-235 (1997)], in which the bacterial expression, purification and enzymatic activity of the HCV NS5B RdRp derived from a patient with chronic sporadic hepatitis is described.

Purified NS5B protein prepared according to Example 2 was incubated in a 10 μL standard reaction mixture consisting of 20 mM HEPES, pH 7.5, 3 mM $MgCl_2$, 1 mM dithiothreitol, 400U/mL RNasin (Gibco/BRL), 0.5 mM each of UTP, ATP and CTP, 0.1 μM [^{32}P]GTP and 0.03 μg pOF1213 RNA at 30°C for 60 minutes. The reaction is terminated by the addition of cold trichloroacetic acid (TCA) and sodium pyrophosphate. The TCA-precipitable

radioactivity was then quantified to determine the extent of RdRp activity.

In addition to the NS5B proteins derived from clones 4, 14, 21, 11, 16 and 20 of the invention, the NS5B gene of the sequence derived from the genotype 1a consensus sequence infectious clone as described by Kolykhalov *et al.* (GenBank Accession Number AF009606) was expressed and purified in the same manner as the proteins of the invention. Table 3 presents a direct comparison of the enzymatic activities of the NS5B proteins of the invention and the consensus sequence protein of Kolykhalov.

Table 3

Comparison of RdRp Activities of Various HCV NS5B Proteins

Protein	specific activity*	amino acid changes relative to clone 4
clone 4 (SEQ ID NO: 2)	9.3	--
clone 14 (SEQ ID NO: 8)	1.7	2
clone 21 (SEQ ID NO: 9)	1.9	5
clone 11 (SEQ ID NO: 10)	0.2	3
clone 16 (SEQ ID NO: 11)	0.1	4
clone 20 (SEQ ID NO: 12)	<0.1	3
AF009606	0.9	71

* pmole NTP incorporated per μg NS5B protein per 60 minutes under standard reaction conditions.

Surprisingly, a considerable range of enzymatic activities was observed among these distinct NS5B proteins. The NS5B protein derived from clone 4 possessed the greatest activity in this comparison, while the activities of clone 14 and 21 were moderate, those of clone 11 and 16 were minimal and that of clone 20 was undetectable. The activity of the genotype 1a consensus NS5B protein of Kolykhalov was substantially less than that clone 4 of the invention. Lohmann *et al.*

reported a specific activity of their genotype 1b NS5B protein on a heteropolymeric RNA template to be 1.7 pmol/μg/120 minutes, similar to the activity reported here for the Kolykhalov enzyme, but also significantly below that of the clone 4 enzyme.

This example demonstrates a range of functional activities among cloned and expressed NS5B sequences, from very active to inactive. In this analysis of 6 NS5B proteins of the invention, there are three instances (clones 11, 16 and 20) of NS5B proteins that are poorly functional or non-functional and three instances (clones 4, 14 and 21) of NS5B proteins that are functional.

It is further revealed here that surprisingly very few changes in amino acid sequence can be sufficient to dramatically alter the enzymatic activity of the NS5B protein. While sequence changes within known or predicted active sites or conserved sequence motifs of an enzyme might be expected to effect an enzyme's activity, most amino acid changes among the 6 NS5B sequences of this comparison lie outside consensus motifs of RdRp enzyme identified by Koonin [J. Gen. Virol. 72:2197-2206 (1991)] and Poch *et al.* [EMBO J. 8:3867-3874 (1989)]. The two exceptions to this are as follows. Clone 16 possesses a single residue change within motif III and VI, and clone 21 possesses a single conservative amino acid change within motif VI (Figure 1). This notwithstanding, the effect of these amino acid changes, and of the others among the NS5B sequences of this example, on the RdRp activity of the NS5B protein is surprising and could not have been

anticipated from prior art.

The data presented in this example also demonstrate that the NS5B RdRp activity derived from clone 4 has dramatically superior activity over that of the NS5B derived from the infectious clone identified by GenBank Accession Number AF009606. This result indicates that the mere generation of a consensus sequence of an NS5B protein does not necessarily provide an NS5B protein with optimal functionality.

EXAMPLE 4

HCV NS5B Sequence Alignments

The NS5B sequences of the invention are distinct from NS5B sequences previously characterized as functional. This is demonstrated by the sequence alignments presented in Figures 3, 4 and 5.

The sequence disclosed by Kolykhalov *et al.* (GenBank Accession Number AF009606) represents a consensus sequence of known NS5B sequences derived from genotype 1a HCV. Figure 3 shows an alignment of the amino acid sequence of clone 4 (SEQ ID NO: 2) of the invention with that of this genotype 1a consensus sequence. There are 71 amino acid differences between these two sequences (excluding the initiator methionine residue in clone 4). This demonstrates that the clone 4 NS5B sequence of the invention is distinct from the genotype 1a consensus sequence. Alignment of this genotype 1a consensus sequence with all other sequences of the invention similarly demonstrates the distinct

nature of the sequences of the invention. Also included in the alignment in Figure 3 is the genotype 1a sequence disclosed in PCT WO 97/12033, the NS5B protein from which was demonstrated to possess RdRp activity. Inspection of this alignment reveals 72 amino acid differences between the sequences, again demonstrating the unique nature of the sequences of the invention.

The sequences of the invention are also distinct from a genotype 1b consensus sequence. A genotype 1b consensus sequence was generated based on alignment of 29 genotype 1b NS5B sequences found in the GenBank database. Comparison of this consensus sequence with that of the clone 4 (SEQ ID NO: 2) of the invention shows that these sequences are distinct (Figure 4). After the initiator methionine (at position -1), the clone 4 sequence differs from this genotype 1b consensus sequence by 13 amino acids. Alignment of this genotype 1b consensus sequence with all other sequences of the invention similarly demonstrates the distinct nature of the sequences of the invention. Also included in the alignment in Figure 4 are the genotype 1b sequences disclosed by Behrens *et al.* (GenBank Accession Number M58335) and by Lohmann *et al.* (GenBank Accession Number Z97730). The NS5B proteins encoded by these sequences have been previously demonstrated to possess RdRp activity. Clone 4 is again distinct in that it possesses 17 amino acid changes relative to M58335 and 12 changes relative to Z97730.

The sequences of the invention are further distinct from a genotype 1b sequence derived from an infectious clone [Yanagi *et al.* Virology 244:161-172 (1998);

GenBank Accession Number AF054247]. As shown in alignment presented in Figure 5, the clone 4 sequence (SEQ ID NO: 2) differs by 25 amino acid residues, excluding the methionine residue at position Đ1 in clone 4.

These alignments reveal the unique nature of sequences of the invention relative to these genotype 1b sequences.

Therefore, based on comparisons with both genotype 1a and genotype 1b consensus sequences, the NS5B sequences of the invention do not represent a consensus sequence. Thus, based on the consensus sequence method for deducing sequence functionality, it would not be apparent that clone 4 is a functional RdRp or a functionally superior RdRp. Moreover, based on comparisons with NS5B sequences previously demonstrated to have functional RdRp activity, the sequence of clone 4 and of the other sequences of the invention are distinct from these RdRp sequences, and again, would not *a priori* be assumed to possess RdRp activity, and further not suspected of possessing superior RdRp activity.

EXAMPLE 5

HCV NS5B Proteins with Improved Activity

Based on sequence comparisons among NS5B sequences of the invention, a number of modified NS5B sequences were constructed in which particular amino acids were changed. These modified sequences were generated by

standard site-specific mutagenesis procedures [Picard *et al.* Nucleic Acid Res. 22:2587-2591 (1994)]. Modified genes were engineered in the baculovirus expression system as described in Example 1, purified according to Example 2 and evaluated for RdRp activity according to Example 4.

The sequence of clone 20 differs from that of clone 4 at three amino acid residues. The presence of a valine (V) residue at position -1 of the clone 20 NS5B protein instead of an initiator methionine (M) suggests that the NS5B protein produced from clone 20 is initiated at the methionine at position 2, and further, that protein produced from this methionine residue is inactive. The V change in clone 20 is likely an artifact of molecular cloning procedures (mistaken sequence in a primer oligonucleotide). The V residue position -1 in clone 20 was changed to M, yielding clone 20(V-1M), identified by SEQ ID NO: 13, which now differs from the sequence of clone 20 by one amino acid (at residue -1) and from clone 4 by two amino acids (at residues 177 and 543). The RdRp activity of the purified NS5B protein derived from clone 20(V-1M) was dramatically increased over that of clone 20 and significantly improved over that of clone 4 (Table 4).

Additional residue changes among the NS5B sequences of the invention can lead to useful and functional proteins. For example, when the alanine (A) residue at position 75 in clone 4 is changed to valine (V), the residue found at this position in clone 14, the resulting clone, clone (A75V), shows considerable RdRp activity (Table 4). In another example, the asparagine

(N) residue in clone 4 at position 177 was changed to aspartic acid (D), the residue found at this position in all other clones of the invention, yielding clone 4(N177D). This clone showed very good RdRp activity (Table 4). In a third example, when the serine (S) at position 543 in clone 4 is changed to proline (P), as is found in clone 20, yielding clone 4(S543P), a functional RdRp is generated (Table 4).

Table 4

RdRp Activities of Various HCV NS5B Proteins

Protein#		specific activity *
clone 20	(SEQ ID NO: 12)	< 0.1
clone 20(V-1M)	(SEQ ID NO: 13)	23.5
clone 4	(SEQ ID NO: 2)	9.3
clone 4(A75V)	(SEQ ID NO: 54)	3.5
clone 4(N177D)	(SEQ ID NO: 55)	7.2
clone 4(S543P)	(SEQ ID NO: 56)	3.5

* pmole NTP incorporated per μg NS5B protein per 60 minutes under standard reaction conditions.

\# All of the changes in table 4 may also be made to the other NS5B encoding amino acid sequences set forth herein i.e. SEQ ID NOS: 8, 9, 10, 11, 12, 13. and thus are contemplated to be within the scope of the invention.

This example demonstrates that amino acid residue changes may be made in the sequence of the NS5B protein and that such changes may maintain the functional activity of the resulting protein. Moreover, the amino acid substitutions may be either conservative or non-conservative in nature.

Additional changes of the amino acid residues among the NS5B sequences of the invention in various combinations are contemplated to be within the scope of the invention.

EXAMPLE 6

Novel HCV NS5B Proteins with Improved Activity

Another strategy for investigating functional NS5B sequences involves the introduction of unique amino acid substitutions at positions where amino acid residues are conserved in known NS5B sequences. For example, all known NS5B sequences have as the amino terminal amino acid (position 1) a serine residue. Several modified versions of clone 4 were constructed in which this serine was substituted with either alanine [clone 4(S1A)], glycine [clone 4(S1G)] or threonine [clone 4(S1T)]. Surprisingly, all of the NS5B proteins derived from these novel sequences exhibited substantial RdRp activity. However, substitution at this position with tyrosine (Y) [clone 4(S1Y)] yielded an inactive NS5B protein (Table 5).

Similarly, all known NS5B sequences have a methionine residue at position 2. Modified versions of clone 4 were constructed in which this methionine was substituted with either alanine [clone 4(M2A)], leucine [clone 4(M2L)] or threonine [clone 4(M2T)]. Surprisingly, all of these novel NS5B proteins exhibited substantial RdRp activity.

In a final example, addition of an amino acid residue prior to the first residue of the mature NS5B protein (serine at position 1), as in clone 4(MAS), in which an alanine residue is inserted before the serine residue, yields a substantially functional NS5B RdRp (Table 5).

These findings, in which amino acid substitutions are made at residues for which there is no naturally occurring variation and that result in the generation of functional, as well as dramatically improved functional, NS5B proteins, were totally unanticipated.

Moreover, these data demonstrate that the nature of amino acid residue changes capable of giving rise to functional variants of NS5B is not limited to conservative amino acid substitutions. For example, while a substitution of threonine for serine at position 1 can be viewed as a conservative amino acid substitution, the substitution with glycine represents a dramatic change in the chemical nature of the amino acid side-chain. Similarly, substitution of the methionine residue at position 2 with leucine represents a conservative change. However, replacement with threonine, a clear non-conservative change, yields a functional and functionally improved NS5B protein. Based on these discoveries, it is therefore fully contemplated here that additional changes of the amino acid residues among the NS5B sequences of the invention, both conservative and non-conservative in nature, at various residue positions and in various combinations will yield useful NS5B compositions and are therefore within the scope of the invention.

Table 5

RdRp Activities of Various HCV NS5B Proteins

Protein#		specific activity *
clone 4	(SEQ ID NO: 2)	9.3
clone 4(S1G)	(SEQ ID NO: 57)	23.0
clone 4(S1A)	(SEQ ID NO: 58)	3.5
clone 4(S1T)	(SEQ ID NO: 59)	3.8
clone 4(S1Y)	(SEQ ID NO: 60)	<0.1
clone 4(M2A)	(SEQ ID NO: 61)	14.0
clone 4(M2L)	(SEQ ID NO: 62)	13.0
clone 4(M2T)	(SEQ ID NO: 64)	22.3
clone 4(MAS)	(SEQ ID NO: 65)	15.0

* pmole NTP incorporated per μg NS5B protein per 60 minutes under standard reaction conditions.

\# while changes to SEQ ID NO2 are exemplified herein, the recited alterations may also be made to the other NS5B amino acid sequences, i.e. SEQ ID NOS: 8, 9, 10, 11, 12, 13. Thus, sequences so altered are also contemplated to be within the scope of the invention

While Examples 5 and 6 demonstrate that various changes in the amino acid sequence of SEQ ID NO: 2 may be made to advantage, it is fully contemplated that the residue changes exemplified here will also have similar utility, either singly or in various combinations, in other NS5B sequences, including sequences of the invention (e.g., 8, 9, 10, 11 and 13, those of Tables 4 and 5), any natural allelic variants, mutants and derivatives of these and other NS5B sequences. That is to say, the changes described in Examples 5 and 6, when introduced into any NS5B sequence, will likely provide a functional or functionally improved NS5B, and are fully contemplated to be within the scope of the present invention.

Several examples of nucleic acid sequences encoding variant NS5B protein sequences useful in the practice of the invention comprise sequences provided in Table 6.

Table 6

Nucleic acid sequences encoding variant NS5B Proteins*

Parental SEQ ID	Codon Position/Change
SEQ ID NO: 1	A codon encoding alanine at position 75 to any codon encoding valine
SEQ ID NO: 1	A codon encoding asparagine at position 177 to any codon encoding aspartic acid
SEQ ID NO: 1	A codon encoding serine at position 543 to any codon encoding proline
SEQ ID NO: 1	A codon encoding serine at position 1 to any codon encoding glycine
SEQ ID NO: 1	A codon encoding serine at position 1 to any codon encoding alanine
SEQ ID NO: 1	A codon encoding serine at position 1 to any codon encoding threonine
SEQ ID NO: 1	A codon encoding serine at position 1 to any codon encoding tyrosine
SEQ ID NO: 1	A codon encoding methionine at position 2 to any codon encoding alanine
SEQ ID NO: 1	A codon encoding methionine at position 2 to any codon encoding leucine
SEQ ID NO: 1	A codon encoding methionine at position 2 to any codon encoding threonine
SEQ ID NO: 1	Insertion of any codon encoding alanine at -1

* While SEQ ID NO: 1 is exemplified in this table, the alterations recited may also be introduced into other NS5B encoding nucleic acid sequences set forth herein, i.e., SEQ ID NOS: 3, 4, 5, 6 and 7. Such altered nucleic acid sequences are also contemplated for use in the present invention.

The examples of Table 6 demonstrate various beneficial changes in the nucleic acid sequence of SEQ ID NO: 1. In view of these results, it is expected that the changes exemplified in SEQ ID NO:1 will also have similar utility, either singly or in various combinations, in other NS5B sequences, including sequences of the invention (e.g., 3, 4, 5, 6 and 7,),

any natural allelic variants, mutants and derivatives of these and other NS5B sequences. That is to say, the changes described in this Table 6, when introduced into any NS5B sequence, will likely provide a functional or functionally improved NS5B, and are thus included in the scope of the present invention.

EXAMPLE 7

Utility of HCV NS5B Protein for Discovery of Antiviral Compositions

The discovery of novel inhibitors of viral polymerases and related proteins often times requires the screening of large numbers of chemical compounds or mixtures of chemical compounds. Thus, an assay for polymerase activity that is capable of high volume screening, in other words, a high throughput assay, is desirable. There are a variety of assay methodologies well known to the trained artisan that allow the efficient screening of large numbers of samples [see, for example, Cole, JL, in Meth. Enzymology 275:310-328 (1996)], and may utilize any number of activity detection and measurement technologies including, but not limited to, radiometric, colorimetric, fluorogenic, or chemiluminescent, any one of which may be suitable in the case of the HCV NS5B RdRp activity.

In one approach, a high throughput assay of the RdRp activity of the functional HCV NS5B proteins of the invention is provided to enable the screening of large numbers of chemicals or other potential inhibitors for

activity against the enzyme. The assay is formatted in 96-well microplates and measures polymerase activity on an RNA template-primer by the incorporation of radiolabeled NTP into trichloroacetic acid (TCA)-precipitable RNA product. Radioactivity may be quantified by either direct scintillation spectrometry or phosphorimaging technology. A phosphorimage of assay results obtained with the NS5B protein of Example 4 [clone 4] for one screening plate is presented in Fig. 6. The first (1) and last (12) columns of the plate contain activity and background controls and a titration of a reference inhibitor compound. Thus, wells A1, B1, A12 and B12 show the activity of the enzyme in the absence of any test compound (100% of the expected activity). In columns 1 and 12, wells C through F, a compound that was discovered to inhibit the HCV RdRp activity by use of the methods of the invention is included in decreasing concentrations. Wells H1 and H12 lack the HCV enzyme and illustrate the background (0% RdRp activity) in the assay. In this example, the remaining 80 wells contain a collection of small organic compounds that were tested for their ability to affect the RdRp activity of the HCV NS5B protein. Thus, it can be seen that the material in wells G2, D10 and F10 represent potent inhibitors of the HCV RdRp activity.

This example demonstrates that HCV NS5B proteins may be used to advantage to identify and assess agents or materials that may affect the HCV RdRp and HCV replication.

In summary, the disclosure of the present invention demonstrates that simple knowledge of an HCV NS5B

sequence is insufficient to allow one to conclude, or even reliably predict that such sequence is a functional sequence or has functional utility. Additionally, an HCV NS5B consensus sequence is not the only, and not necessarily the optimal, functional NS5B protein encoding sequence. The present invention provides novel NS5B sequences with unanticipated functionality. Additionally, the present invention demonstrates that particular sequence changes in HCV NS5B can lead to unanticipated and significantly improved functionality. In light of the foregoing, the NS5B sequences of the invention represent novel sequences with demonstrated functionality and utility and unique improvements over prior art. Finally, the HCV NS5B sequences of the invention have broad utility in research, diagnostic, therapeutic and pharmaceutical applications.

While certain embodiments of the present invention have been described and/or exemplified above, various other embodiments will be apparent to those skilled in the art from the foregoing disclosure. The present invention is, therefore, not limited to the particular embodiments described and/or exemplified, but is capable of considerable modification without departure from the scope of the appended claims.

What is claimed is:

1. A nucleic acid molecule encoding a hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 1 and natural allelic variants, mutants and derivatives thereof.

2. A nucleic acid molecule encoding a hepatitis C virus NS5B protein having a sequence of SEQ ID NO. 2.

3. A nucleic acid molecule encoding a hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 3 and natural allelic variants, mutants and derivatives thereof.

4. A nucleic acid molecule encoding a hepatitis C virus NS5B protein having a sequence of SEQ ID NO: 8

5. A nucleic acid molecule encoding a hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 4 and natural allelic variants, mutants and derivatives thereof.

6. A nucleic acid molecule encoding a hepatitis C virus NS5B protein having the sequence of SEQ ID NO: 9.

7. A nucleic acid molecule encoding a hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 5 and natural allelic variants, mutants and derivatives thereof.

8. A nucleic acid molecule encoding a hepatitis C virus NS5B protein having the sequence of SEQ ID NO: 10.

9. A nucleic acid molecule encoding a hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 6 and natural allelic variants, mutants and derivatives thereof.

10. A nucleic acid molecule encoding a hepatitis C virus NS5B protein having the sequence of SEQ ID NO: 11.

11. A nucleic acid molecule encoding a hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 7 and natural allelic variants, mutants and derivatives thereof.

12. A nucleic acid molecule encoding a hepatitis C virus NS5B protein having the sequence of SEQ ID NO: 12.

13. A nucleic acid molecule encoding a hepatitis C virus protein, said molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7 and natural allelic variants, mutants and derivatives thereof.

14. A nucleic acid as claimed in claim 13, said nucleic acid molecule being modified as set forth in Table 6.

15. A nucleic acid molecule encoding a hepatitis C

virus NS5B protein, said protein having an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 and SEQ ID NO: 13.

16. A hepatitis C virus NS5B protein having a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 and SEQ ID NO: 13.

17. An HCV NS5B protein as claimed in claim 16, said protein being modified as set forth in Tables 4 and 5.

18. A vector comprising a nucleic acid sequence encoding a hepatitis C virus NS5B protein, said protein consisting essentially of an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, and SEQ ID NO: 13.

19. A vector as claimed in claim 18, said vector comprising a nucleic acid encoding a hepatitis C NS5B protein, said amino acid sequence of said protein being altered as set forth in Tables 4 and 5.

20. A hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 2.

21. A hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 8.

22. A hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 9.

23. A hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 10.

24. A hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 11.

25. A hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 12.

26. A hepatitis C virus NS5B protein comprising the sequence of SEQ ID NO: 13.

27. A hepatitis C virus NS5B protein having an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, and SEQ ID NO: 13, wherein the amino acid residues at positions +1 and +2 may be any amino acid.

28. A hepatitis C virus NS5B protein as claimed in claim 27, wherein said amino acid sequence is altered as set forth in Tables 4 and 5.

29. A nucleic acid molecule encoding a hepatitis C virus NS5B protein in which the amino acid at position 177 of the mature polypeptide is asparagine.

30. A hepatitis C virus NS5B protein in which the

amino acid at position 177 of the mature polypeptide is asparagine.

31. A nucleic acid molecule as claimed in claim 1, wherein the codon encoding an amino acid at position 75 is changed from alanine to valine.

32. A hepatitis C virus NS5B protein as claimed in claim 20, wherein the amino acid at position 75 is changed from alanine to valine.

33. A nucleic acid molecule as claimed in claim 1, wherein the codon encoding an amino acid at position 177 is changed from asparagine to aspartic acid.

34. A hepatitis C virus NS5B protein as claimed in claim 20, wherein the amino acid at position 177 is changed from asparagine to aspartic acid.

35. A nucleic acid molecule as claimed in claim 1, wherein a codon encoding an amino acid at position 534 is changed from serine to proline.

36. A hepatitis C virus NS5B protein as claimed in claim 20, wherein the amino acid at position 534 is changed from serine to proline.

37. A nucleic acid molecule encoding a HCV NS5B protein having a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7, wherein the

codons encoding amino acids at positions -1, -2, and -3 may encode any amino acid.

38. A nucleic acid molecule as claimed in claim 13, further comprising a tag sequence.

39. A nucleic acid molecule as claimed in claim 38, said tag sequence being selected from the group consisting of homopolymeric nucleic acid sequences, polyhistidine, flag epitope, c-myc epitope, transmembrane epitope of the influenza A virus hemaglutinin protein,, protein A, cellulose binding domain, calmodulin binding protein, maltose binding protein, chitin binding domain, glutathione S-transferase, or biotin.

40. A nucleic acid molecule as claimed in claim 14, further comprising a tag sequence.

41. A nucleic acid molecule as claimed in claim 40, said tag sequence being selected from the group consisting of homopolymeric nucleic acid sequences, polyhistidine, flag epitope, c-myc epitope, transmembrane epitope of the influenza A virus hemaglutinin protein,, protein A, cellulose binding domain, calmodulin binding protein, maltose binding protein, chitin binding domain, glutathione S-transferase, or biotin.

42. A hepatitis C virus NS5B protein selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ

ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 and SEQ ID NO: 13 and a modified NS5B protein as set forth in Tables 4 and 5, further comprising a protein tag sequence, said tag sequence being selected from the group consisting of homopolymeric nucleic acid sequences, polyhistidine, flag epitope, c-myc epitope, transmembrane epitope of the influenza A virus hemaglutinin protein,, protein A, cellulose binding domain, calmodulin binding protein, maltose binding protein, chitin binding domain, glutathione S-transferase, or biotin.

43. A nucleic acid molecule encoding a hepatitis C virus protein having a nucleic acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7, natural allelic variants, mutants and derivatives of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7 and nucleic acids modified as set forth in Table 6, wherein an amino acid codon for alanine precedes the first amino acid codon of the mature NS5B protein.

44. A hepatitis C virus NS5B protein having a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 and SEQ ID NO: 13 and a modified NS5B protein as set forth in Tables 4 and 5, wherein an alanine residue precedes the first amino acid of the mature NS5B protein.

45. A method for assaying a test compound, for a

modulating activity against hepatitis C virus comprising:

i) providing an enzymatically active hepatitis C virus NS5B protein comprising a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 13;

ii) contacting said protein with a test compound suspected of modulating hepatitis C virus NS5B activity; and

iii) measuring modulation of said hepatitis C virus NS5B activity by said test compound.

46. A method as claimed in claim 45, wherein said HCV NS5B protein sequence is modified as set forth in Tables 4 and 5.

47. A method as claimed in claim 45, wherein the amino acids at positions +1 and +2 in said HCV NS5B protein may be any amino acid.

48. A method as claimed in claim 45, wherein amino acids at positions -1, -2 and -3 of said HCV NS5B protein may be any amino acid.

49. A method as claimed in claim 45, wherein said HCV NS5B protein further comprises a tag sequence.

50. A method for assaying a compound, for antagonistic activity against hepatitis C virus comprising:

i) providing an enzymatically active hepatitis C virus NS5B protein comprising a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 13;

ii) contacting said protein with a test compound suspected of antagonizing hepatitis C virus NS5B activity; and

iii) measuring antagonism of said hepatitis C virus NS5B activity by said test compound.

51. A method as claimed in claim 50, wherein said HCV NS5B protein sequence is modified as set forth in Tables 4 and 5.

52. A method as claimed in claim 50, wherein the amino acids at positions +1 and +2 in said HCV NS5B protein may be any amino acid.

53. A method as claimed in claim 50, wherein said HCV NS5B protein further comprises a tag sequence.

54. A method as claimed in claim 50, wherein amino acids at positions -1, -2 and -3 of said HCV NS5B protein may be any amino acid.

55. A method for assaying a test compound, for agonistic activity against hepatitis C virus comprising:

i) providing an enzymatically active hepatitis C virus NS5B protein comprising a sequence selected from the group consisting of SEQ ID NO: 2, SEQ

ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 13;

ii) contacting said protein with a test compound suspected of agonizing hepatitis C virus NS5B activity; and

iii) measuring agonism of said hepatitis C virus NS5B activity by said test compound.

56. A method as claimed in claim 55, wherein said HCV NS5B protein sequence is modified as set forth in Tables 4 and 5.

57. A method as claimed in claim 55, wherein the amino acids at positions +1 and +2 in said HCV NS5B protein may be any amino acid.

58. A method as claimed in claim 55, wherein amino acids at positions -1, -2 and -3 of said HCV NS5B protein may be any amino acid.

59. A method as claimed in claim 55, wherein said HCV NS5B protein further comprises a tag sequence.

60. A method for assaying a test compound for interaction with the hepatitis C virus NS5B protein sequence comprising:

i) providing a hepatitis C virus NS5B polypeptide consisting essentially of a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, and SEQ ID NO: 13;

ii) contacting said protein or peptide with a test compound suspected of interacting with the hepatitis C virus NS5B protein or peptide; and

iii) measuring the interaction of said test compound or material with said hepatitis C virus NS5B polypeptide.

61. A method as claimed in claim 60, wherein said HCV NS5B protein sequence is modified as set forth in Tables 4 and 5.

62. A method as claimed in claim 60, wherein the amino acids at positions +1 and +2 in said HCV NS5B protein may be any amino acid.

63. A method as claimed in claim 60, wherein amino acids at positions -1, -2 and -3 of said HCV NS5B protein may be any amino acid.

64. A method as claimed in claim 60, wherein said HCV NS5B protein further comprises a tag sequence.

65. A method for assaying a test compound, for interaction with a hepatitis C virus NS5B nucleic acid sequence comprising:

i) providing a hepatitis C virus NS5B nucleic acid consisting essentially of a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, and SEQ ID NO: 7;

ii) contacting said NS5B nucleic acid with a test compound suspected of interacting with the

hepatitis C virus NS5B NS5B nucleic acid; and

iii) measuring the interaction of said test compound with said hepatitis C virus NS5B NS5B nucleic acid.

66. A method as claimed in claim 65, wherein said HCV NS5B nucleic acid sequence is modified as set forth in Table 6.

67. A method for detecting the presence of HCV in a biological sample, comprising nucleic acid amplification that effectively amplifies a selected nucleotide sequence from said virus and detection of said selected sequence; said nucleic acid having a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, and SEQ ID NO: 7.

68. A method as claimed in claim 67, said nucleic acid sequence being altered as set forth in Table 6.

69. A method as claimed in claim 67, wherein the nucleic acid amplification method is performed employing a series of nested primers.

70. A method as claimed in claim 67, wherein said sequence is detected using an oligonucleotide comprising a sequence set forth in Table 1.

71. A method for detecting immunological interactions between viral polypeptides and antibodies

directed toward hepatitis C virus in a biological sample, said method comprising isolating said antibodies using an NS5B amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, and SEQ ID NO: 13.

72. A method as claimed in claim 71, wherein said HCV NS5B amino acid sequence is modified as set forth in Tables 4 and 5.

73. A method as claimed in claim 71, wherein the amino acids at positions +1 and +2 in said HCV NS5B protein may be any amino acid.

74. A method as claimed in claim 71, wherein amino acids at positions -1, -2 and -3 of said HCV NS5B protein can be any amino acid.

75. An antibody having affinity for a hepatitis C polypeptide having a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, and SEQ ID NO: 13.

76. An antibody as claimed in claim 75, wherein said HCV NS5B protein sequence is modified as set forth in Tables 4 and 5.

77. An antibody as claimed in claim 75, wherein the amino acids at positions +1 and +2 in said HCV NS5B

protein may be any amino acid.

78. An antibody as claimed in claim 75, wherein amino acids at positions -1, -2 and -3 of said HCV NS5B protein may be any amino acid.

79. An antibody having affinity for a polypeptide comprising a sequence set forth in Table 2.

80. A nucleic acid molecule encoding a hepatitis C virus protein having a nucleic acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7, natural allelic variants, mutants and derivatives of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7 and nucleic acids modified as set forth in Table 6, which encodes a viral antigen.

81. A method for introducing a nucleic acid as claimed in claim 80 into a host, wherein said nucleic acid is delivered to said host via a process selected from the group consisting of transformation, transfection, transduction, transgenetics, surgically and by physical bombardment with nucleic acid-coated particles.

82. A viral antigen comprising an amino acid sequence
selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 and SEQ ID NO: 13 and a modified NS5B

protein as set forth in Tables 4 and 5

83. A viral antigen as claimed in claim 82, wherein the amino acids at positions +1 and +2 in said HCV NS5B protein may be any amino acid.

84. A viral antigen as claimed in claim 82, wherein amino acids at positions -1, -2 and -3 of said HCV NS5B protein may be any amino acid.

85. A viral antigen as claimed in claim 82, further comprising a protein tag sequence.

86. A method of raising an immune response in a mammalian subject comprising administering to said subject a viral antigen encoded by the nucleic acid of claim 80.

87. A method of raising an immune response in a mammalian subject comprising administering to said subject a viral antigen according to claim 82.

88. A method of amplifying hepatitis C virus nucleic acids employing primers comprising sequences selected from the oligonucleotide sequences set forth in Table 1.

89. A host cell comprising a hepatitis C virus nucleic acid as claimed 13.

90. A host cell comprising a hepatitis C virus

nucleic acid, said nucleic acid being altered as set forth in Table 6.

91. A host cell comprising a nucleic acid encoding a hepatitis C NS5B protein, said nucleic acid being selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7, natural allelic variants, mutants and derivatives of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7 and nucleic acids modified as set forth in Table 6, said host cell being selected from the group of cells consisting of bacterial cells, fungal cells, insect cells, mammalian cells, and plant cells.

92. A cell line comprising a nucleic acid encoding a hepatitis C NS5B protein, said nucleic acid being selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7, natural allelic variants, mutants and derivatives of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7 and nucleic acids modified as set forth in Table 6, said cell line being selected from the group consisting of a hepatocyte cell line, Chang liver cells, Hela cells, U937 cells, HepG2 cells, MT-4 cells and clonal cells generated from said cell lines.

93. A cell line according to claim 92, wherein said cell line expresses a functional NS5B protein.

94. A cell line according to claim 93 further

comprising a reporter system wherein a reporting signal which from said reporter system is dependent on the polymerase activity of a functional HCV NS5B protein in said cells.

95. A cell line according to claim 94, wherein said reporter system includes a nucleic acid comprising the antisense strand of an RNA molecule encoding a functional reporter molecule, said NS5B polymerase acting on said antisense molecule to generate a translatable mRNA sequence in said host cell, said translated reporter molecule sequence being capable of producing an increase in signal which is indicative of the presence of a functional NS5B polymerase.

96. A cell line according to claim 95 wherein said functional mRNA molecule encodes a reporter molecule selected from a group consisting of luciferase, β-galactosidase, alkaline phosphatase or fluorescent green protein host cell is a hepatocyte.

97. A method for assessing the functionality of a hepatitis C NS5B protein, comprising:

i) providing a cell line as claimed in claim 92;

ii) transforming said cell line with a nucleic acid encoding an antisense molecule, the complement of which encodes a functional reporter molecule, said NS5B polymerase acting on said antisense molecule to generate a translatable mRNA sequence in said host cell; and

iii) assessing said cells for an increase in signal produced by said translated reporter molecule sequence, said increase indicating the presence of a functional NS5B protein.

98. A method as claimed in claim 97, wherein said functional reporter molecule is selected from a group consisting of luciferase, β-galactosidase, alkaline phosphatase or fluorescent green protein.

99. A host cell as claimed in claim 89, wherein said host cell is a hepatocyte.

100. A method for assaying a test compound for antagonist activity against hepatitis C virus comprising:

i) providing a cell line according to claim 93;

ii) contacting said cell line with a test compound suspected of having antagonistic hepatitis C virus NS5B activity; and

iii) measuring the antagonistic effect, if any, of said test compound on hepatitis C virus NS5B activity.

101. A method for assaying a test compound for agonist activity against hepatitis C virus comprising:

i) providing a cell line according to claim 93;

ii) contacting said cell line with a test compound suspected of having agonistic hepatitis C virus

NS5B activity; and

iii) measuring the agonistic effect, if any, of said test compound on hepatitis C virus NS5B activity.

102. A method of preparing a hepatitis C virus NS5B protein comprising a sequence selected from a group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, and SEQ ID NO: 13 comprising:

i) expressing said protein in a biological system;

ii) solubilizing said protein;

iii) extracting said protein and

iii) enriching for said NS5B protein.

103. A method as claimed in claim 102, wherein said hepatitis C virus protein has been altered as set forth in Tables 4 and 5.

104. A method as claimed in claim 102, wherein the amino acids at positions +1 and +2 in said HCV NS5B protein can be any amino acid.

105. A method as claimed in claim 102, wherein amino acids at positions -1, -2 and -3 of said HCV NS5B protein may be any amino acid.

106. A method as claimed in claims 102, wherein said HCV NS5B further comprising a protein tag sequence according to claim 40.

107. A method as claimed in claim 102, wherein said NS5B protein is enriched by a process selected from the group consisting of cell disruption, cell fractionation, partitioning, reverse micelle partitioning, aqueous two-phase extraction, precipitation, chromatography, ion exchange chromatography, chelation chromatography, affinity chromatography, immunoaffinity chromatography, high pressure liquid chromatography, hydrophobic interaction chromatography, centrifugation, membrane filtration, gel filtration, immunoprecipitation, electrophoresis, isoelectric focusing, and isotachophoresis.

108. A method as claimed in claim 102, wherein said biological system is selected from the group consisting of bacterial cells, fungal cells, insect cells, mammalian cells, and plant cells.

109. A method as claimed in claim 102, wherein said biological system is cell free.

110. A method for isolating an HCV NS5B protein, comprising:

i) contacting an extract of a biological system containing the NS5B protein with a solid matrix possessing a diethylaminoethyl (DEAE) moiety as a functional group in a salt-containing buffer solution;

ii) removing said NS5B protein from said solid matrix and further contacting said protein with a solid matrix comprising heparin as a functional group;

iii) exposing said heparin matrix containing bound NS5B protein to a buffer solution of continuously increasing levels of salt;

iv) collecting NS5B protein-containing fractions and contacting said collected fractions with a solid matrix containing Cibacron Blue as a functional group;

v) exposing said Cibacron Blue matrix containing bound NS5B protein to a buffered solution containing increasing levels of salt; and

vi) collecting said enriched NS5B containing fractions.

111. A method of generating an infectious viral vector, comprising incorporating a hepatitis C virus NS5B nucleic acid selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, and SEQ ID NO: 7 into said viral vector.

112. A method as claimed in claim 111, wherein said nucleic acid sequence is altered as set forth in Table 6.

113. A method of generating an infectious viral vector comprising substituting an HCV NS5B homologous gene sequence derived from any HCV sequence with a hepatitis C virus NS5B nucleic acid selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, and SEQ ID NO: 7 in said viral vector.

114. A method as claimed in claim 113, wherein said nucleic acid sequence is altered as set forth in Table 6.

115. A host animal comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7, natural allelic variants, mutants and derivatives of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7 and nucleic acids modified as set forth in Table 6.

116. A method of isolating antibodies to HCV NS5B protein comprising screening a human or murine antibody library for reactivity to NS5B proteins having a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 and SEQ ID NO: 13 and a modified NS5B protein as set forth in Tables 4 and 5, and selecting a clone from said library which expresses the reactive antibody and isolating said clone.

117. A method for propagating HCV in an *in vitro* system comprising culturing a cell line in which functional HCV NS5B sequences are provided, said sequences being selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7, natural allelic variants, mutants and derivatives of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 5, SEQ ID NO:6, and SEQ ID NO: 7 and nucleic acids modified as set forth in Table 6, under

conditions that facilitate the replication of HCV.

118. A method for infecting a cell culture with hepatitis C virus comprising administering a viral vector which is generated as claimed in claim 111.

119. A method for infecting an animal with hepatitis C virus comprising administering a viral vector which is generated as claimed in claim 111.

120. A method of propagating hepatitis C virus *in vitro* comprising:

i) contacting a cell line with a functional HCV NS5B protein consisting essentially of a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, and SEQ ID NO: 13; and

ii) said cell line being further contacted with an HCV nucleic acid.

121. A method as claimed in claim 120, wherein said NS5B protein has been modified as set forth in Tables 4 and 5

122. A method as claimed in claim 120, wherein the amino acid residues at positions +1 and +2 in said NS5B protein may be any amino acid.

123. A method according to claim 120 wherein said NS5B nucleic acid is provided within a hepatitis C virus replicon.

124. A method of propagating hepatitis C virus in a living host, comprising:

i) delivering a nucleic acid encoding a hepatitis C virus NS5B protein having a sequence selected from the group consisting SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, and SEQ ID NO: 13 to a living host, said nucleic acid encoding a functional NS5B protein and being operably linked to an expression operon; and

ii) contacting said living host with an HCV molecule selected from the group consisting of HCV particles, HCV nucleic acids, and HCV replicons.

125. A method as claimed in claim 124, wherein said NS5B encoding nucleic acid is altered as set forth in Table 6.

126. A method as claimed in claim 124, wherein said nucleic acid is delivered to said host via a process selected from the group consisting of transformation, transfection, transduction, transgenetics, surgically or by physical bombardment with nucleic acid-coated particles.

127. A method according to claim 124 wherein said NS5B encoding nucleic acid of step i) is provided within a replicon.

128. A kit for detecting the presence of HCV nucleic acids in a biological sample, comprising:

i) oligonucleotides comprising a sequence set

forth in Table I, said oligonucleotides being hybridizable to said HCV encoding nucleic acid;

ii) reaction buffer; and

iii) an instruction sheet.

129. A kit as claimed in claim 128, wherein said oligonucleotide contains a tag.

130. A kit for detecting the presence of HCV in a biological sample, comprising:

i) antibodies immunologically specific for HCV proteins;

ii) a solid support with immobilized HCV antigens as a positive control; and

iii) an instruction sheet.

131. A kit as claimed in claim 130, wherein said antibody contains a tag.

```
MSMSYTWTGAL ITPCAAEESK LPINALSNSL LRHHNMVYAT TSRSASQRQK 50  Clone  4; SEQ ID NO: 2
----------- ---------- ---------- ---------- ----------      Clone 14; SEQ ID NO: 8
----------- ---------- ---------- ---------- ----------      Clone 21; SEQ ID NO: 9
----------- ---------- ---------- ---------- ----------      Clone 11; SEQ ID NO: 10
-------R--- ---------- ---------- ---------- ----------      Clone 16; SEQ ID NO: 11
V---------- ---------- ---------- ---------- ----------      Clone 20; SEQ ID NO: 12

KVTFDRLQVL DDHYRDVLKE MKAKASTVKA KLLSVEEACK LTPPHSARSK 100
---------- ---------- ----V----- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------

                                                  I
                                             |----------|
FGYGAKDVRN LSSKAVNHIH SVWKDLLEDT ETPIDTTIMA KNEVFCVQPE 150
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------

          II                                      III
  |--------------------|                     |------------
KGGRKPARLI VYPDLGVRVC EKMALYNVVS TLPQAVMGSS YGFQYSPGQR 200
---------- ---------- ------D--- ---------- ----------
---------- ---------- ------D-I- ---------- ----------
---------- ---------- ------D--- ---------- ----------
---------- ---------- ------D--- ---------- ----H-----
---------- ---------- ------D--- ---------- ----------

                IV/A
--------|  |-----------------|
VEFLVNAWKS KKCPMGFAYD TRCFDSTVTE NDIRVEESIY QCCDLAPEAR 250
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------P ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------

                                           V/B
                                  |--------------------------
QAIRSLTERL YIGGPLTNSK GQSCGYRRCR ASGVLTTSCG NTLTCYLKAS 300
---------- ---------- ---------- ---------- ----------
---------- ---------- -----H---- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------


            VI/C
---|        |-----------|
AACRAAKLQD CTMLVCGDDL VVICESAGTQ EDAASLRVFT EAMTRYSAPP 350
---------- ---------- ---------- ---------- ----------
---------- ---------- I--------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------P ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------


          VII/D                VIII
        |------------|       |--------------|
GDPPQPEYDL ELITSCSSNV SVAHDASGKR VYYLTRDPTT PLARAAWETA 400
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
```

Fig. 1A

```
RHTPVNSWLG NIIMYAPTLW ARMILMTHFF SILLAQEQLE KALDCEIYGA 450
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- -T-------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------

HYSIEPLDLP QIIQRLHGLS APSLHSYSPG EINRVASCLR KLGVPPLRVW 500
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------

RHRARSVRAK LLSQGGRAAT CGKYLFNWAV RTKLKLTPIP AASQLDLSGW 550
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- --P-------

FVAGYSGGDI YHSLSRARPR WFMWCLLLLS VGVGIYLLPN R          591   Clone  4; SEQ ID NO: 2
---------- ---------- ---------- ---------- -                 Clone 14; SEQ ID NO: 8
---------- ---------C ---------- ---------- -                 Clone 21; SEQ ID NO: 9
---------- ---------- ---------- ---------- -                 Clone 11; SEQ ID NO: 10
---------- ---------- ---------- ---------- -                 Clone 16; SEQ ID NO: 11
---------- ---------- ---------- ---------- -                 Clone 20; SEQ ID NO: 12
```

Fig. 1B

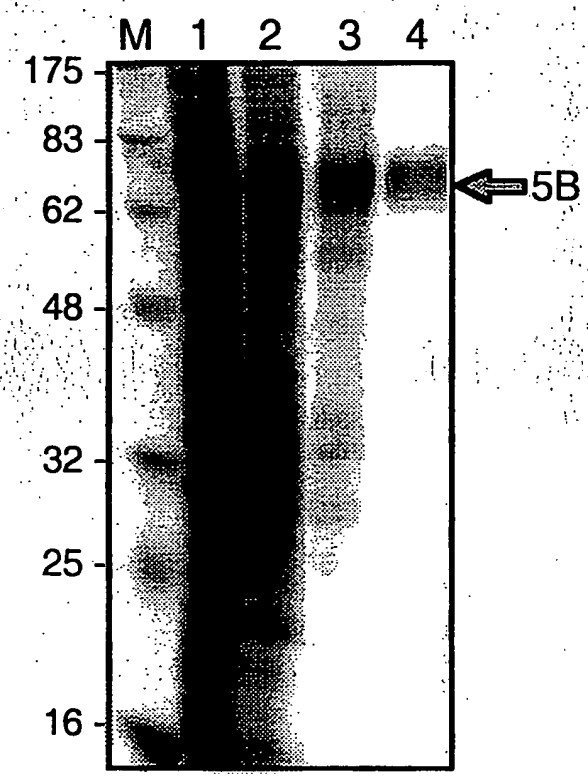

Fig. 2

```
 SMSYSWTGAL VTPCAAEEQK LPINALSNSL LRHHNLVYST TSRSACQRQK  50  AF009606
MA---------- ---------- ---------- ---------- ----------      WO 97/12033
 M----T----- I-------S- ---------- -----M--A- -----S----      Clone 4; SEQ ID NO: 2

 KVTFDRLQVL DSHYQDVLKE VKAAASKVKA NLLSVEEACS LTPPHSAKSK 100
 ---------- ---------- ---------- ---------- ----------
 ---------- -D--R----- M--K--T--- K---------K -------R--

 FGYGAKDVRC HARKAVAHIN SVWKDLLEDS VTPIDTTIMA KNEVFCVQPE 150
 ---------- ------T--- ---------N ---------- ----------
 ---------N LSS---N--H ---------T E--------- ----------

 KGGRKPARLI VFPDLGVRVC EKMALYDVVS KLPLAVMGSS YGFQYSPGQR 200
 ---------- ---------- ---------T ---------- ----------
 ---------- -Y--------- ------N--S T--Q------ ----------

 VEFLVQAWKS KRTPMGFSYD TRCFDSTVTE SDIRTEEAIY QCCDLDPQAR 250
 ---------- ---------- ---------- ---------- ----------
 -----N----- --C----A-- ---------- N---V--S-- -----A-E--

 VAIKSLTERL YVGGPLTNSR GENCGYRRCR ASGVLTTSCG NTLTCYIKAR 300
 ---------- ---------- ---------- ---------- ----------
 Q--R------ -I-------K -QS------- ---------- ------L--S

 AACRAAGLQD CTMLVCGDDL VVICESAGVQ EDAASLRAFT EAMTRYSAPP 350
 ---------- ---------- ---------- ---------- ----------
 ------K--- ---------- --------T- -------V-- ----------

 GDPPQPEYDL ELITSCSSNV SVAHDGAGKR VYYLTRDPTT PLARAAWETA 400
 ---------- ---------- ---------- ---------- ----------
 ---------- ---------- -----AS--- ---------- ----------

 RHTPVNSWLG NIIMFAPTLW ARMILMTHFF SVLIARDQLE QALNCEIYGA 450
 ---------- ---------- ---------- ---------- ---D------
 ---------- ----Y----- ---------- -I-L-Q---- K--D------

 CYSIEPLDLP PIIQRLHGLS AFSLHSYSPG EINRVAACLR KLGVPPLRAW 500
 ---------- ---------- ---------- ---------- ----------
 H--------- Q--------- ---------- ------S--- --------V-

 RHRARSVRAR LLSRGGRAAI CGKYLFNWAV RTKLKLTPIA AAGRLDLSGW 550
 ---------- --A------- ---------- ---------- ---Q------
 ---------K ---Q-----T ---------- ---------P --SQ------

 FTAGYSGGDI YHSVSHARPR WFWFCLLLLA AGVGIYLLPN R        591  AF009606
 ---------- ---------- -I-------- ---------- -             WO 97/12033
 -V--------- ---L-R---- --MW-----S V--------- -             Clone 4; SEQ ID NO: 2
```

Fig. 3

```
SMSYTWTGAL ITPCAAEESK LPINALSNSL LRHHNMVYAT TSRSASLRQK  50   1b consensus
M---------- ---------- ---------- ---------- -----G----       M58335
M---------- ---------- ---------- ---------- ----------       Z97730
M---------- ---------- ---------- ---------- ------Q---       Clone 4 (SEQ ID NO:2)

KVTFDRLQVL DDHYRDVLKE MKAKASTVKA KLLSVEEACK LTPPHSAKSK 100
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- -------R--

FGYGAKDVRN LSSRAVNHIR SVWKDLLEDT ETPIDTTIMA KNEVFCVQPE 150
---------- ---K-----H ---------- V--------- ----------
---------- ---K------ ---------- ---------- ----------
---------- ---K-----H ---------- ---------- ----------

KGGRKPARLI VFPDLGVRVC EKMALYDVVS TLPQAVMGSS YGFQYSPGQR 200
---------- ---------- ---------- ----V----- ----------
---------- ---------- ---------- ---------- ----------
---------- -Y-------- ------N--- ---------- ----------

VEFLVNAWKS KKCPMGFSYD TRCFDSTVTE NDIRVEESIY QCCDLAPEAR 250
------T--- --N------- ---------- ---------- ----------
------T--- ---------- ---------- ---------- ----------
---------- -------A-- ---------- ---------- ----------

QAIRSLTERL YIGGPLTNSK GQNCGYRRCR ASGVLTTSCG NTLTCYLKAS 300
---K------ ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- --S------- ---------- ----------

AACRAAKLQD CTMLVNGDDL VVICESAGTQ EDAASLRVFT EAMTRYSAPP 350
---------- ---------- ---------- ---------- ----------
---------- -----C---- ---------- -------A-- ----------
---------- -----C---- ---------- ---------- ----------

GDPPQPEYDL ELITSCSSNV SVAHDASGKR VYYLTRDPTT PLARAAWETA 400
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------

RHTPVNSWLG NIIMYAPTLW ARMILMTHFF SILLAQEQLE KALDCQIYGA 450
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- -----E----

CYSIEPLDLP QIIERLHGLS APSLHSYSPG EINRVASCLR KLGVPPLRVW 500
---------- ---------- ---------- ---------- ----------
---------- ---Q------ ---------- ---------- ----------
H--------- ---Q------ ---------- ---------- ----------

RHRARSVRAK LLSQGGRAAT CGKYLFNWAV RTKLKLTPIP AASQLDLSGW 550
---------R ---------- ---------- K--------- ---R------
---------R ---------- ---------- ---------- ----------
---------- ---------- ---------- ---------- ----------

FVAGYSGGDI YHSLSRARPR WFMLCLLLLS VGVGIYLLPN R         591   1b consensus
---------- ---------- ---------- ---------- -               M58335
---------- ---------- ---W------ ---------- -               Z97730
---------- ---------- ---W------ ---------- -               Clone 4 (SEQ ID NO:2)
```

Fig. 4

```
MSMSYTWTGAL ITPCAAEESK LPINALSNSL LRHHNMVYAT TSRSASQRQK  50  Clone 4; SEQ ID NO:2
 ---------- ---------- ----P----- ---------- ------L---      AF054247

KVTFDRLQVL DDHYRDVLKE MKAKASTVKA KLLSVEEACK LTPPHSARSK 100
 ---------- ---------- ---------- ----I----- -------K--

                                                    I
                                              |----------|
PGYGAKDVRN LSSKAVNHIH SVWKDLLEDT ETPIDTTIMA KNEVFCVQPE 150
 ---------- ---R-----R ---------- ---------- -S--------

            II                                      III
  |-------------------|                    |------------
KGGRKPARLI VYPDLGVRVC EKMALYNVVS TLPQAVMGSS YGFQYSPGQR 200
 ---------- -F--------- ------D--- ---------- -------K--

                  IV/A
 --------|  |----------------|
VEFLVNAWKS KKCPMGFAYD TRCFDSTVTE NDIRVEESIY QCCDLAPEAR 250
 ------T--- -------S-- ---------- S--------- ----------

                                                 V/B
                                |-------------------------
QAIRSLTERL YIGGPLTNSK GQSCGYRRCR ASGVLTTSCG NTLTCYLKAS 300
 ---------- ---------- --N------- ---------- ---------T

               VI/C
 ---|        |-----------|
AACRAAKLQD CTMLVCGDDL VVICESAGTQ EDAASLRVFT EAMTRYSAPP 350
 ---------- -----N---- ---------- ----A--A-- ----------

              VII/D                  VIII
          |------------|       |--------------|
GDPPQPEYDL ELITSCSSNV SVAHDASGKR VYYLTRDPTT PLARAAWETA 400
 ---------- ---------- ---------- ---------- ----------

RHTPVNSWLG NIIMYAPTLW ARMILMTHFF SILLAQEQLE KALDCEIYGA 450
 ----I----- ---------- ---------- ---------- -----Q----

HYSIEPLDLP QIIQRLHGLS AFSLHSYSPG EINRVASCLR KLGVPPLRVW 500
 C--------- ---Q------ --T------- ---------- ----------

RHRARSVRAK LLSQGGRAAT CGKYLFNWAV RTKLKLTPIP AASQLDLSGW 550
 ---------- ---------- --R------- ---------- ----------

FVAGYSGGDI YHSLSRARPR WFMWCLLLLS VGVGIYLLPN R          591  Clone 4; SEQ ID NO:2
 ---------- ---------- --PL------ ---------- -               AF054247
```

Fig. 5

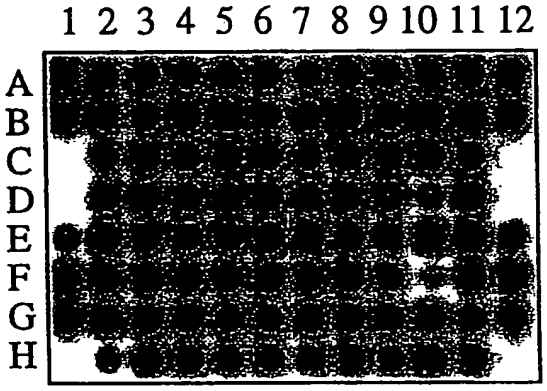

Fig. 6

SEQ ID NO: 1 (Clone 4)

ATGTCAATGTCCTACACATGGACAGGCGCTCTGATTACACCATGCGCTGC 50
GGAGGAGAGCAAGCTGCCCATCAATGCGCTGAGCAACTCTTTGCTGCGCC 100
ACCATAACATGGTCTATGCCACAACATCCCGCAGCGCAAGCCAGCGGCAG 150
AAGAAGGTAACTTTTGACAGGCTGCAAGTCCTGGATGACCACTACCGGGA 200
CGTGCTCAAGGAGATGAAGGCGAAGGCGTCCACAGTCAAGGCTAAACTTC 250
TATCCGTGGAAGAAGCCTGTAAGCTGACGCCCCCACATTCAGCCAGATCC 300
AAATTTGGCTATGGGGCGAAGGACGTCCGGAACCTATCCAGCAAGGCCGT 350
TAACCACATCCACTCCGTGTGGAAGGACTTGCTGGAAGACACTGAGACAC 400
CAATTGACACCACCATCATGGCAAAAAATGAGGTTTTCTGCGTTCAACCA 450
GAGAAAGGAGGCCGCAAGCCAGCTCGCCTAATCGTGTACCCAGACTTGGG 500
GGTTCGCGTGTGCGAGAAAATGGCCCTCTACAACGTGGTCTCCACCCTTC 550
CTCAGGCTGTGATGGGCTCCTCATACGGATTCCAGTACTCTCCTGGACAG 600
CGGGTCGAGTTCCTGGTGAATGCCTGGAAGTCAAAGAAATGCCCTATGGG 650
CTTCGCATATGACACCCGCTGTTTCGACTCAACGGTCACCGAGAATGACA 700
TCCGTGTTGAGGAGTCAATTTACCAATGTTGTGACTTGGCCCCCGAAGCC 750
AGACAGGCCATAAGGTCGCTTACAGAGCGGCTTTATATCGGGGGTCCCCT 800
GACTAACTCAAAAGGGCAGAGCTGCGGTTATCGCCGGTGCCGTGCGAGCG 850
GCGTACTGACGACTAGCTGCGGTAATACCCTCACATGTTACTTGAAGGCC 900
TCTGCAGCCTGTCGAGCTGCAAAGCTCCAGGACTGCACGATGCTTGTGTG 950
CGGAGACGACCTCGTCGTTATCTGTGAAAGCGCGGGAACCCAAGAGGACG 1000
CGGCGAGCCTACGAGTCTTCACGGAGGCTATGACTAGGTACTCTGCCCCC 1050
CCTGGGGACCCGCCCCAACCAGAATACGACTTGGAGCTAATAACATCATG 1100
CTCCTCCAACGTGTCGGTCGCGCACGATGCATCTGGCAAAAGGGTATACT 1150
ACCTCACCCGTGACCCCACCACCCCTCTTGCGCGGGCTGCGTGGGAGACA 1200
GCTAGACACACTCCAGTTAACTCCTGGCTAGGCAACATCATCATGTATGC 1250
GCCCACTCTGTGGGCAAGGATGATTCTGATGACTCACTTCTTCTCCATCC 1300
TTCTAGCTCAGGAGCAACTTGAAAAAGCCCTAGATTGTGAGATCTACGGG 1350
GCCCACTACTCCATTGAGCCACTTGACCTACCTCAGATCATTCAACGACT 1400
CCATGGTCTTAGCGCGTTTTCACTCCACAGTTACTCTCCAGGTGAAATCA 1450
ATAGGGTGGCTTCATGCCTCAGGAAGCTTGGGGTACCACCCTTGCGAGTC 1500
TGGAGACATCGGGCCAGAAGCGTCCGCGCTAAGTTACTGTCCCAGGGGGG 1550
GAGGGCTGCCACTTGTGGCAAGTACCTCTTCAACTGGGCAGTAAGGACCA 1600
AGCTTAAACTCACTCCAATTCCGGCTGCGTCCCAGTTGGACTTATCCGGC 1650
TGGTTCGTTGCTGGTTACAGCGGGGGAGACATATATCACAGCCTGTCTCG 1700
TGCCCGACCCCGCTGGTTCATGTGGTGCCTACTCCTACTTTCTGTAGGGG 1750
TAGGCATCTACCTGCTCCCGAACCGATGA 1779

Fig. 7

SEQ ID NO: 4 (Clone 21)

ATGTCAATGTCCTACACATGGACAGGCGCTCTGATTACACCATGCGCTGC 50
GGAGGAGAGCAAGCTGCCCATCAATGCGCTGAGCAACTCTTTGCTGCGCC 100
ACCATAACATGGTCTATGCCACAACATCCCGCAGCGCAAGCCAGCGGCAG 150
AAGAAGGTAACTTTTGACAGGCTGCAAGTCCTGGATGACCACTACCGGGA 200
CGTGCTCAAGGAGATGAAGGCGAAGGCGTCCACAGTCAAGGCTAAACTTC 250
TATCCGTGGAAGAAGCCTGTAAGCTGACGCCCCCACATTCAGCCAGATCC 300
AAATTTGGCTATGGGGCGAAGGACGTCCGGAACCTATCCAGCAAGGCCGT 350
TAACCACATCCACTCCGTGTGGAAGGACTTGCTGGAAGACACTGAGACAC 400
CAATTGACACCACCATCATGGCAAAAAATGAGGTTTTCTGCGTTCAACCA 450
GAGAAAGGAGGCCGCAAGCCAGCTCGCCTAATCGTGTACCCAGACTTGGG 500
GGTTCGCGTGTGCGAGAAAATGGCCCTCTACGACGTGATCTCCACCCTTC 550
CTCAGGCTGTGATGGGCTCCTCATACGGATTCCAGTACTCTCCTGGACAG 600
CGGGTCGAGTTCCTGGTGAATGCCTGGAAGTCAAAGAAATGCCCTATGGG 650
CTTCGCATATGACACCCGCTGTTTCGACTCGACGGTCACCGAGAATGACA 700
TCCGTGTTGAGGAGTCAATTTACCAATGTTGTGACTTGGCCCCCGAAGCC 750
AGACAGGCCATAAGGTCGCTTACAGAGCGGCTTTATATCGGGGGTCCCCT 800
GACTAACTCAAAAGGGCAGAGCTGCGGTCATCGCCGGTGCCGTGCGAGCG 850
GCGTACTGACGACTAGCTGCGGTAATACCCTCACATGTTACTTGAAGGCC 900
TCTGCAGCCTGTCGAGCTGCAAAGCTCCAGGACTGCACGATGCTTGTGTG 950
CGGAGACGACCTCATCGTTATCTGTGAAAGCGCGGGAACCCAAGAGGACG 1000
CGGCGAGCCTACGAGTCTTCACGGAGGCTATGACTAGGTACTCTGCCCCC 1050
CCTGGGGACCCGCCCCAACCAGAATACGACTTGGAGCTAATAACATCATG 1100
CTCCTCCAACGTGTCGGTCGCGCACGATGCATCTGGCAAAAGGGTATACT 1150
ACCTCACCCGTGACCCCACCACCCCTCTTGCGCGGGCTGCGTGGGAGACA 1200
GCTAGACACACTCCAGTTAACTCCTGGCTAGGCAACATCATCATGTATGC 1250
GCCCACTCTGTGGGCAAGGATGATTCTGATGACTCACTTCTTCTCCATCC 1300
TTCTAGCTCAGGAGCAACTTGAAAAAGCCCTAGATTGTGAGATCTACGGG 1350
GCCCACTACTCCATTGAGCCACTTGACCTACCTCAGATCATTCAACGACT 1400
CCATGGTCTTAGCGCGTTTTCACTCCACAGTTACTCTCCAGGTGAAATCA 1450
ATAGGGTGGCTTCATGCCTCAGGAAGCTTGGGGTACCACCCTTGCGAGTC 1500
TGGAGACATCGGGCCAGAAGCGTCCGCGCTAAGTTACTGTCCCAGGGGGG 1550
GAGGGCTGCCACTTGTGGCAAGTACCTCTTCAACTGGGCAGTAAGGACCA 1600
AGCTTAAACTCACTCCAATTCCGGCTGCGTCCCAGTTGGACTTATCCGGC 1650
TGGTTCGTTGCTGGTTACAGCGGGGGAGACATATACCACAGCCTGTCTCG 1700
TGCCCGACCCTGCTGGTTCATGTGGTGCCTACTCCTACTTTCTGTAGGGG 1750
TAGGCATCTACCTGCTCCCCAACCGATGA 1779

Fig. 9

SEQ ID NO: 6 (Clone 16)

ATGTCAATGTCCTACACATGGAGAGGCGCTCTGATTACACCATGCGCTGC 50
GGAGGAGAGCAAGCTGCCCATCAATGCGCTGAGCAACTCTTTGCTGCGCC 100
ACCATAACATGGTCTATGCCACAACATCCCGCAGCGCAAGCCAGCGGCAG 150
AAGAAGGTAACTTTTGACAGGCTGCAAGTCCTGGATGACCACTACCGGGA 200
CGTGCTCAAGGAGATGAAGGCGAAGGCGTCCACAGTCAAGGCCAAACTTC 250
TATCCGTGGAAGAAGCCTGTAAGCTGACGCCCCCACATTCAGCCAGATCC 300
AAATTTGGCTATGGGGCGAAGGACGTCCGGAACCTATCCAGCAAGGCCGT 350
TAACCACATCCACTCCGTGTGGAAGGACTTGCTGGAAGACACTGAGACAC 400
CAATTGACACCACCATCATGGCAAAAAATGAGGTCTTCTGCGTTCAACCA 450
GAGAAAGGAGGCCGCAAGCCAGCTCGCCTAATCGTGTACCCAGACTTGGG 500
GGTTCGCGTGTGCGAGAAAATGGCCCTCTACGACGTGGTCTCCACCCTTC 550
CTCAGGCTGTGATGGGCTCCTCATACGGATTCCAGCACTCTCCTGGACAG 600
CGGGTCGAGTTCCTGGTGAATGCCTGGAAGTCAAAGAAATGCCCTATGGG 650
CTTCGCATATGACACCCGCTGTTTCGACTCAACGGTCACCGAGAATGACA 700
TCCGTGTTGAGGAGTCAATTTACCAATGTTGTGACTTGGCCCCCGAAGCC 750
AGACAGGCCATAAGGTCGCTTACAGAGCGGCTTTATATCGGGGGTCCCCT 800
GACTAACTCAAAAGGGCAGAGCTGCGGTTATCGCCGGTGCCGTGCGAGCG 850
GCGTACTGACGACTAGCTGCGGTAATACCCTCACATGTTACTTGAAGGCC 900
TCTGCAGCCTGTCGAGCTGCAAAGCTCCAGGACTGCACGATGCTTGTGTG 950
CGGAGACGACTTCGTCGTTATCTGTGAAAGCGCGGGAACCCAAGAGGACG 1000
CGGCGAGCCTACGAGTCTTCACGGAGGCTATGACTAGGTACTCTGCCCCC 1050
CCTGGGGACCCGCCCCAACCAGAATACGACTTGGAGCTAATAACATCATG 1100
CTCCTCCAACGTGTCGGTCGCGCACGATGCATCTGGCAAAAGGGTATACT 1150
ACCTCACCCGTGACCCCACCACCCCTCTTGCGCGGGCTGCGTGGGAGACA 1200
GCTAGACACACTCCAGTTAACTCCTGGCTAGGCAACATCATCATGTATGC 1250
GCCCACTCTGTGGGCAAGGATGATTCTGATGACTCACTTCTTCTCCATCC 1300
TTCTAGCTCAGGAGCAACTTGAAAAAGCCCTAGATTGTGAGATCTACGGG 1350
GCCCACTACTCCATTGAGCCACTTGACCTACCTCAGATCATTCAACGACT 1400
CCATGGTCTTAGCGCGTTTTCACTCCACAGTTACTCTCCAGGTGAAATCA 1450
ATAGGGTGGCTTCATGCCTCAGGAAGCTTGGGGTACCACCCTTGCGAGTC 1500
TGGAGACATCGGGCCAGAAGCGTCCGCGCTAAGTTACTGTCCCAGGGGGG 1550
GAGGGCTGCCACTTGTGGCAAGTACCTCTTCAACTGGGCAGTAAGGACCA 1600
AGCTTAAACTCACTCCAATTCCGGCTGCGTCCCAGTTGGACTTATCCGGC 1650
TGGTTCGTTGCTGGTTACAGCGGGGGAGACATATATCACAGCCTGTCTCG 1700
TGCCCGACCCCGCTGGTTCATGTGGTGCCTACTCCTACTTTCTGTAGGGG 1750
TAGGCATCTACCTGCTCCCCAACCGATGA 1779

Fig. 11

SEQ ID NO: 7 (Clone 20)

GTGTCAATGTCCTACACATGGACAGGCGCTCTGATTACACCATGCGCTGC 50
GGAGGAGAGCAAGCTGCCCATCAATGCGCTGAGCAACTCTTTGCTGCGCC 100
ACCATAACATGGTCTATGCCACAACATCCCGCAGCGCAAGCCAGCGGCAG 150
AAGAAGGTAACTTTTGACAGGCTGCAAGTCCTGGATGACCACTACCGGGA 200
CGTGCTCAAGGAGATGAAGGCGAAGGCGTCCACAGTCAAGGCTAAACTTC 250
TATCCGTGGAAGAAGCCTGTAAGCTGACGCCCCCACATTCAGCCAGATCC 300
AAATTTGGCTATGGGGCGAAGGACGTCCGGAACCTATCCAGCAAGGCCGT 350
TAACCACATCCACTCCGTGTGGAAGGACTTGCTGGAAGACACTGAGACAC 400
CAATTGACACCACCATCATGGCAAAAAATGAGGTTTTCTGCGTTCAACCA 450
GAGAAAGGAGGCCGCAAGCCAGCTCGCCTAATCGTGTACCCAGACTTGGG 500
GGTTCGCGTGTGCGAGAAAATGGCCCTCTACGACGTGGTCTCCACCCTTC 550
CTCAGGCTGTGATGGGCTCCTCATACGGATTCCAGTACTCTCCTGGACAG 600
CGGGTCGAGTTCCTGGTGAATGCCTGGAAGTCAAAGAAATGCCCTATGGG 650
CTTCGCATATGACACTCGCTGTTTCGACTCAACGGTCACCGAGAATGACA 700
TCCGTGTTGAGGAGTCAATTTACCAATGTTGTGACTTGGCCCCCGAAGCC 750
AGACAGGCCATAAGGTCGCTTACAGAGCGGCTTTATATCGGGGGTCCCCT 800
GACTAACTCAAAAGGGCAGAGCTGCGGTTATCGCCGGTGCCGTGCGAGCG 850
GCGTACTGACGACTAGCTGCGGTAATACCCTCACATGTTACTTGAAGGCC 900
TCTGCAGCCTGTCGAGCTGCAAAGCTCCAGGACTGCACGATGCTTGTGTG 950
CGGAGACGACCTCGTCGTTATCTGTGAAAGCGCGGGAACCCAAGAGGACG 1000
CGGCGAGCCTACGAGTCTTCACGGAGGCTATGACTAGGTACTCTGCCCCC 1050
CCTGGGGACCCGCCCCAACCAGAATACGACTTGGAGCTAATAACATCATG 1100
CTCCTCCAACGTGTCGGTCGCGCACGATGCATCTGGCAAAAGGGTATACT 1150
ACCTCACCCGTGACCCCACCACCCCTCTTGCGCGGGCTGCGTGGGAGACA 1200
GCTAGACACACTCCAGTTAACTCCTGGCTAGGCAACATCATCATGTATGC 1250
GCCCACTCTGTGGGCAAGGATGATTCTGATGACTCACTTCTTCTCCATCC 1300
TTCTAGCTCAGGAGCAACTTGAAAAAGCCCTAGATTGTGAGATCTACGGG 1350
GCCCACTACTCCATTGAGCCACTTGACCTACCTCAGATCATTCAACGACT 1400
CCATGGTCTTAGCGCGTTTTCACTCCACAGTTACTCTCCAGGTGAAATCA 1450
ATAGGGTGGCTTCATGCCTCAGGAAGCTTGGGGTACCACCCTTGCGAGTC 1500
TGGAGACATCGGGCCAGAAGCGTCCGCGCTAAGTTACTGTCCCAGGGGGG 1550
GAGGGCTGCCACTTGTGGCAAGTACCTCTTCAACTGGGCAGTAAGGACCA 1600
AGCTTAAACTCACTCCAATTCCGGCTGCGCCCCAGTTGGACTTATCCGGC 1650
TGGTTCGTTGCTGGTTACAGCGGGGGAGACATATATCACAGCCTGTCTCG 1700
TGCCCGACCCCGCTGGTTCATGTGGTGCCTACTCCTACTTTCTGTAGGGG 1750
TAGGCATCTACCTGCTCCCCAACCGATGA 1779

Fig. 12

SEQ ID NO: 13 (Clone 20(V-1M))

-1 M	SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASQRQK	50
	KVTFDRLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSK	100
	FGYGAKDVRNLSSKAVNHIHSVWKDLLEDTETPIDTTIMAKNEVFCVQPE	150
	KGGRKPARLIVYPDLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQR	200
	VEFLVNAWKSKKCPMGFAYDTRCFDSTVTENDIRVEESIYQCCDLAPEAR	250
	QAIRSLTERLYIGGPLTNSKGQSCGYRRCRASGVLTTSCGNTLTCYLKAS	300
	AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPP	350
	GDPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETA	400
	RHTPVNSWLGNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCEIYGA	450
	HYSIEPLDLPQIIQRLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVW	500
	RHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIPAAPQLDLSGW	550
	FVAGYSGGDIYHSLSRARPRWFMWCLLLLSVGVGIYLLPNR	591

Fig. 13

SEQUENCE LISTING

<110> Viropharma, Incorporated

<120> Hepatitis C Virus NS5B Compositions and Methods of Use Thereof

<130> HEPPOL

<150> 60/080,509
<151> 1998-04-02

<150> 60/090,356
<151> 1998-06-23

<160> 53

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 1779
<212> DNA
<213> Hepatitis C Virus

<400> 1
atgtcaatgt cctacacatg gacaggcgct ctgattacac catgcgctgc ggaggagagc 60
aagctgccca tcaatgcgct gagcaactct ttgctgcgcc accataacat ggtctatgcc 120
acaacatccc gcagcgcaag ccagcggcag aagaaggtaa cttttgacag gctgcaagtc 180
ctggatgacc actaccggga cgtgctcaag gagatgaagg cgaaggcgtc cacagtcaag 240
gctaaacttc tatccgtgga agaagcctgt aagctgacgc ccccacattc agccagatcc 300
aaatttggct atggggcgaa ggacgtccgg aacctatcca gcaaggccgt taaccacatc 360
cactccgtgt ggaaggactt gctggaagac actgagacac caattgacac caccatcatg 420
gcaaaaaatg aggttttctg cgttcaacca gagaaaggag gccgcaagcc agctcgccta 480
atcgtgtacc cagacttggg ggttcgcgtg tgcgagaaaa tggccctcta caacgtggtc 540
tccacccttc ctcaggctgt gatgggctcc tcatacggat tccagtactc tcctggacag 600
cgggtcgagt tcctggtgaa tgcctggaag tcaaagaaat gccctatggg cttcgcatat 660
gacacccgct gtttcgactc aacggtcacc gagaatgaca tccgtgttga ggagtcaatt 720
taccaatgtt gtgacttggc ccccgaagcc agacaggcca taaggtcgct tacagagcgg 780
ctttatatcg ggggtccgct gactaactca aaagggcaga gctgcggtta tcgccggtgc 840
cgtgcgagcg gcgtactgac gactagctgc ggtaataccc tcacatgtta cttgaaggcc 900
tctgcagcct gtcgagctgc aaagctccag gactgcacga tgcttgtgtg cggagacgac 960
ctcgtcgtta tctgtgaaag cgcgggaacc caagaggacg cggcgagcct acgagtcttc 1020
acggaggcta tgactaggta ctctgccccc cctggggacc cgccccaacc agaatacgac 1080
ttggagctaa taacatcatg ctcctccaac gtgtcggtcg cgcacgatgc atctggcaaa 1140
agggtatact acctcacccg tgaccccacc acccctcttg cgcgggctgc gtgggagaca 1200
gctagacaca ctccagttaa ctcctggcta ggcaacatca tcatgtatgc gcccactctg 1260
tgggcaagga tgattctgat gactcacttc ttctccatcc ttctagctca ggagcaactt 1320
gaaaaagccc tagattgtga gatctacggg gcccactact ccattgagcc acttgaccta 1380
cctcagatca ttcaacgact ccatggtctt agcgcgtttt cactccacag ttactctcca 1440
ggtgaaatca atagggtggc ttcatgcctc aggaagcttg gggtaccacc cttgcgagtc 1500
tggagacatc gggccagaag cgtccgcgct aagttactgt cccaggggg gagggctgcc 1560
acttgtggca agtacctctt caactgggca gtaaggacca agcttaaact cactccaatt 1620
ccggctgcgt cccagttgga cttatccggc tggttcgttg ctggttacag cgggggagac 1680
atatatcaca gcctgtctcg tgcccgaccc cgctggttca tgtggtgcct actcctactt 1740
tctgtagggg taggcatcta cctgctcccg aaccgatga 1779

<210> 2
<211> 592
<212> PRT

<213> Hepatitis C Virus

<400> 2

```
Met Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala
 1               5                  10                  15
Ala Glu Glu Ser Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu
            20                  25                  30
Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Ser Gln
        35                  40                  45
Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His
    50                  55                  60
Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Ala Ser Thr Val Lys
65                  70                  75                  80
Ala Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro Pro His
                85                  90                  95
Ser Ala Arg Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu
            100                 105                 110
Ser Ser Lys Ala Val Asn His Ile His Ser Val Trp Lys Asp Leu Leu
        115                 120                 125
Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu
    130                 135                 140
Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu
145                 150                 155                 160
Ile Val Tyr Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu
                165                 170                 175
Tyr Asn Val Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser Ser Tyr
            180                 185                 190
Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Ala
        195                 200                 205
Trp Lys Ser Lys Lys Cys Pro Met Gly Phe Ala Tyr Asp Thr Arg Cys
    210                 215                 220
Phe Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile
225                 230                 235                 240
Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Arg Ser
                245                 250                 255
Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly
            260                 265                 270
Gln Ser Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr
        275                 280                 285
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys
    290                 295                 300
Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp
305                 310                 315                 320
Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser
                325                 330                 335
Leu Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly
            340                 345                 350
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser
        355                 360                 365
Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr
    370                 375                 380
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr
385                 390                 395                 400
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr
                405                 410                 415
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser
            420                 425                 430
Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Glu Ile
        435                 440                 445
Tyr Gly Ala His Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile
    450                 455                 460
```

```
Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro
465                 470                 475                 480
Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro
                485                 490                 495
Pro Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Lys Leu
            500                 505                 510
Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn
        515                 520                 525
Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser
    530                 535                 540
Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp
545                 550                 555                 560
Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Met Trp Cys
                565                 570                 575
Leu Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg
            580                 585                 590
```

<210> 3
<211> 1779
<212> DNA
<213> Hepatitis C Virus

<400> 3

```
atgtcaatgt cctacacatg gacaggcgct ctgattacac catgcgctgc ggaggagagc      60
aagctgccca tcaatgcgct gagcaactct ttgctgcgcc accataacat ggtctatgcc     120
acaacatccc gcagcgcaag ccagcggcag aagaaggtaa cttttgacag gctgcaagtc     180
ctggatgacc actaccggga cgtgctcaag gagatgaagg cgaaggtgtc cacagtcaag     240
gctaaacttc tatccgtgga agaagcctgt aagctgacgc ccccacattc agccagatcc     300
aaatttggct atggggcgaa ggacgtccgg aacctatcca gcaaggccgt taaccacatc     360
cactccgtgt ggaaggactt gctggaagac actgagacac caattgacac caccatcatg     420
gcaaaaaatg aggttttctg cgttcaacca gagaaaggag gccgcaagcc agctcgccta     480
atcgtgtacc cagacttggg ggttcgcgtg tgcgagaaaa tggccctcta cgacgtggtc     540
tccacccttc ctcaggctgt gatgggctcc tcatacggat tccagtactc tcctggacag     600
cgggtcgagt tcctggtgaa tgcctggaag tcaaagaaat gccctatggg cttcgcatat     660
gacacccgct gtttcgactc aacggtcacc gagaatgaca tccgtgttga ggagtcaatt     720
taccaatgtt gtgacttggc ccccgaagcc agacaggcca taaggtcgct tacagagcgg     780
ctttatatcg ggggtccccc gactaactca aaagggcaga gctgcggtta tcgccggtgc     840
cgtgcgagcg gcgtactgac gactagctgc ggtaataccc tcacatgtta cttgaaggcc     900
tctgcagcct gtcgagctgc aaagctccag gactgcacga tgcttgtgtg cggagacgac     960
ctcgtcgtta tctgtgaaag cgcgggaacc caagaggacg cggcgagcct acgagtcttc    1020
acggaggcta tgactaggta ctctgccccc cctggggacc cgccccaacc agaatacgac    1080
ttggagctaa taacatcatg ctcctccaac gtgtcggtcg cgcacgatgc atctggcaaa    1140
agggtatact acctcacccg tgaccccacc acccctcttg cgcgggctgc gtgggagaca    1200
gctagacaca ctccagttaa ctcctggcta ggcaacatca tcatgtatgc gcccactctg    1260
tgggcaagga tgattctgat gactcacttc ttctccatcc ttctagctca ggagcaactt    1320
gaaaaagccc tagattgtga gatctacggg gcccactact ccattgagcc acttgaccta    1380
cctcagatca ttcaacgact ccatggtctt agcgcgtttt cactccacag ttactctcca    1440
ggtgaaatca atagggtggc ttcatgcctc aggaagcttg gggtaccacc cttgcgagtc    1500
tggagacatc gggccagaag cgtccgcgct aagttactgt cccaggggggg gagggctgcc    1560
acttgtggca agtacctctt caactgggca gtaaggacca agcttaaact cactccaatt    1620
ccggctgcgt cccagttgga cttatccggc tggttcgttg ctggttacag cgggggagac    1680
atatatcaca gcctgtctcg tgcccgaccc cgctggttca tgtggtgcct actcctactt    1740
tctgtagggg taggcatcta cctgctcccc aaccgatga                           1779
```

<210> 4
<211> 1779
<212> DNA
<213> Hepatitis C Virus

<400> 4

```
atgtcaatgt cctacacatg gacaggcgct ctgattacac catgcgctgc ggaggagagc      60
```

aagctgccca tcaatgcgct gagcaactct ttgctgcgcc accataacat ggtctatgcc 120
acaacatccc gcagcgcaag ccagcggcag aagaaggtaa cttttgacag gctgcaagtc 180
ctggatgacc actaccggga cgtgctcaag gagatgaagg cgaaggcgtc cacagtcaag 240
gctaaacttc tatccgtgga agaagcctgt aagctgacgc ccccacattc agccagatcc 300
aaatttggct atggggcgaa ggacgtccgg aacctatcca gcaaggccgt taaccacatc 360
cactccgtgt ggaaggactt gctggaagac actgagacac caattgacac caccatcatg 420
gcaaaaaatg aggttttctg cgttcaacca gagaaaggag gccgcaagcc agctcgccta 480
atcgtgtacc cagacttggg ggttcgcgtg tgcgagaaaa tggccctcta cgacgtgatc 540
tccacccttc ctcaggctgt gatgggctcc tcatacggat tccagtactc tcctggacag 600
cgggtcgagt tcctggtgaa tgcctggaag tcaaagaaat gccctatggg cttcgcatat 660
gacacccgct gtttcgactc gacggtcacc gagaatgaca tccgtgttga ggagtcaatt 720
taccaatgtt gtgacttggc ccccgaagcc agacaggcca taaggtcgct tacagagcgg 780
ctttatatcg ggggtcccct gactaactca aaagggcaga gctgcggtca tcgccggtgc 840
cgtgcgagcg gcgtactgac gactagctgc ggtaataccc tcacatgtta cttgaaggcc 900
tctgcagcct gtcgagctgc aaagctccag gactgcacga tgcttgtgtg cggagacgac 960
ctcatcgtta tctgtgaaag cgcgggaacc caagaggacg cggcgagcct acgagtcttc 1020
acggaggcta tgactaggta ctctgccccc cctggggacc cgccccaacc agaatacgac 1080
ttggagctaa taacatcatg ctcctccaac gtgtcggtcg cgcacgatgc atctggcaaa 1140
agggtatact acctcacccg tgaccccacc acccctcttg cgcgggctgc gtgggagaca 1200
gctagacaca ctccagttaa ctcctggcta ggcaacatca tcatgtatgc gcccactctg 1260
tgggcaagga tgattctgat gactcacttc ttctccatcc ttctagctca ggagcaactt 1320
gaaaaagccc tagattgtga gatctacggg gcccactact ccattgagcc acttgaccta 1380
cctcagatca ttcaacgact ccatggtctt agcgcgtttt cactccacag ttactctcca 1440
ggtgaaatca atagggtggc ttcatgcctc aggaagcttg gggtaccacc cttgcgagtc 1500
tggagacatc gggccagaag cgtccgcgct aagttactgt cccaggggg gagggctgcc 1560
acttgtggca agtacctctt caactgggca gtaaggacca agcttaaact cactccaatt 1620
ccggctgcgt cccagttgga cttatccggc tggttcgttg ctggttacag cgggggagac 1680
atataccaca gcctgtctcg tgcccgaccc tgctggttca tgtggtgcct actcctactt 1740
tctgtagggg taggcatcta cctgctcccc aaccgatga 1779

<210> 5
<211> 1779
<212> DNA
<213> Hepatitis C Virus

<400> 5
atgtcaatgt cctacacatg gacaggcgct ctgattacac catgcgctgc ggaggagagc 60
aagctgccca tcaatgcgct gagcaactct ttgctgcgcc accataacat ggtctatgcc 120
acaacatccc gcagcgcaag ccagcggcag aagaaggtaa cttttgacag gctgcaagtc 180
ctggatgacc actaccggga cgtgctcaag gagatgaagg cgaaggcgtc cacagtcaag 240
gctaaacttc tatccgtgga agaagcctgt aagctgacgc ccccacattc agccagatcc 300
aaatttggct atggggcgaa ggacgtccgg aacctatcca gcaaggccgt taaccacatc 360
cactccgtgt ggaaggactt gctggaagac actgagacac caattgacac caccatcatg 420
gcaaaaaatg aggttttctg cgttcaacca gagaaaggag gccgcaagcc agctcgccta 480
atcgtgtacc cagacttggg ggttcgcgtg tgcgagaaaa tggccctcta cgacgtggtc 540
tccacccttc ctcaggctgt gatgggctcc tcatacggat tccagtactc ccctggacag 600
cgggtcgagt tcctggtgaa tgcctggaag ccaaagaaat gccctatggg cttcgcatat 660
gacacccgct gtttcgactc aacggtcacc gagaatgaca tccgtgttga ggagtcaatt 720
taccaatgtt gtgacttggc ccccgaagcc agacaggcca taaggtcgct tacagagcgg 780
ctttatatcg ggggtcccct gactaactca aaagggcaga gctgcggtta tcgccggtgc 840
cgtgcgagcg gcgtactgac gactagctgc ggtaataccc tcacatgtta cttgaaggcc 900
tctgcagcct gtcgagctgc aaagctccag gactgcacga tgcttgtgtg cggagacgac 960
ctcgtcgtta tctgtgaaag cgcgggaacc caagaggacg cggcgagcct acgagtcttc 1020
acggaggcta tgactaggta ctctgccccc cctggggacc cgccccaacc agaatacgac 1080
ttggagctaa taacatcatg ctcctccaac gtgtcggtcg cgcacgatgc atctggcaaa 1140
agggtatact acctcacccg tgaccccacc acccctcttg cgcgggctgc gtgggagaca 1200
gctagacaca ctccagttaa ctcctggcta ggcaacacca tcatgtatgc gcccactctg 1260
tgggcaagga tgattctgat gactcacttc ttctccatcc ttctagctca ggagcaactt 1320
gaaaaagccc tagattgtga gatctacggg gcccactact ccattgagcc acttgaccta 1380
cctcagatca ttcaacgact ccatggtctt agcgcgtttt cactccacag ttactctcca 1440
ggtgaaatca atagggtggc ttcatgcctc aggaagcttg gggtaccacc cttgcgagtc 1500

tggagacatc gggccagaag cgtccgcgct aagttactgt cccaggggg gagggctgcc 1560
acttgtggca agtacctctt caactgggca gtaaggacca agcttaaact cactccaatt 1620
ccggctgcgt cccagttgga cttatccggc tggttcgttg ctggttacag cgggggagac 1680
atatatcaca gcctgtctcg tgcccgaccc cgctggttca tgtggtgcct actcctactt 1740
tctgtagggg taggcatcta cctgctcccc aaccgatga 1779

<210> 6
<211> 1779
<212> DNA
<213> Hepatitis C Virus

<400> 6
atgtcaatgt cctacacatg gagaggcgct ctgattacac catgcgctgc ggaggagagc 60
aagctgccca tcaatgcgct gagcaactct ttgctgcgcc accataacat ggtctatgcc 120
acaacatccc gcagcgcaag ccagcggcag aagaaggtaa cttttgacag gctgcaagtc 180
ctggatgacc actaccggga cgtgctcaag gagatgaagg cgaaggcgtc cacagtcaag 240
gccaaacttc tatccgtgga agaagcctgt aagctgacgc ccccacattc agccagatcc 300
aaatttggct atggggcgaa ggacgtccgg aacctatcca gcaaggccgt taaccacatc 360
cactccgtgt ggaaggactt gctggaagac actgagacac caattgacac caccatcatg 420
gcaaaaaatg aggtcttctg cgttcaacca gagaaaggag gccgcaagcc agctcgccta 480
atcgtgtacc cagacttggg ggttcgcgtg tgcgagaaaa tggccctcta cgacgtggtc 540
tccacccttc ctcaggctgt gatgggctcc tcatacggat tccagcactc tcctggacag 600
cgggtcgagt tcctggtgaa tgcctggaag tcaaagaaat gccctatggg cttcgcatat 660
gacacccgct gtttcgactc aacggtcacc gagaatgaca tccgtgttga ggagtcaatt 720
taccaatgtt gtgacttggc ccccgaagcc agacaggcca taaggtcgct tacagagcgg 780
ctttatatcg ggggtccccc gactaactca aaagggcaga gctgcggtta tcgccggtgc 840
cgtgcgagcg gcgtactgac gactagctgc ggtaataccc tcacatgtta cttgaaggcc 900
tctgcagcct gtcgagctgc aaagctccag gactgcacga tgcttgtgtg cggagacgac 960
ttcgtcgtta tctgtgaaag cgcgggaacc caagaggacg cggcgagcct acgagtcttc 1020
acggaggcta tgactaggta ctctgccccc cctggggacc cgccccaacc agaatacgac 1080
ttggagctaa taacatcatg ctcctccaac gtgtcggtcg cgcacgatgc atctggcaaa 1140
agggtatact acctcacccg tgaccccacc acccctcttg cgcgggctgc gtgggagaca 1200
gctagacaca ctccagttaa ctcctggcta ggcaacatca tcatgtatgc gcccactctg 1260
tgggcaagga tgattctgat gactcacttc ttctccatcc ttctagctca ggagcaactt 1320
gaaaaagccc tagattgtga gatctacggg gcccactact ccattgagcc acttgaccta 1380
cctcagatca ttcaacgact ccatggtctt agcgcgtttt cactccacag ttactctcca 1440
ggtgaaatca atagggtggc ttcatgcctc aggaagcttg gggtaccacc cttgcgagtc 1500
tggagacatc gggccagaag cgtccgcgct aagttactgt cccaggggg gagggctgcc 1560
acttgtggca agtacctctt caactgggca gtaaggacca agcttaaact cactccaatt 1620
ccggctgcgt cccagttgga cttatccggc tggttcgttg ctggttacag cgggggagac 1680
atatatcaca gcctgtctcg tgcccgaccc cgctggttca tgtggtgcct actcctactt 1740
tctgtagggg taggcatcta cctgctcccc aaccgatga 1779

<210> 7
<211> 1779
<212> DNA
<213> Hepatitis C Virus

<400> 7
gtgtcaatgt cctacacatg gacaggcgct ctgattacac catgcgctgc ggaggagagc 60
aagctgccca tcaatgcgct gagcaactct ttgctgcgcc accataacat ggtctatgcc 120
acaacatccc gcagcgcaag ccagcggcag aagaaggtaa cttttgacag gctgcaagtc 180
ctggatgacc actaccggga cgtgctcaag gagatgaagg cgaaggcgtc cacagtcaag 240
gctaaacttc tatccgtgga agaagcctgt aagctgacgc ccccacattc agccagatcc 300
aaatttggct atggggcgaa ggacgtccgg aacctatcca gcaaggccgt taaccacatc 360
cactccgtgt ggaaggactt gctggaagac actgagacac caattgacac caccatcatg 420
gcaaaaaatg aggttttctg cgttcaacca gagaaaggag gccgcaagcc agctcgccta 480
atcgtgtacc cagacttggg ggttcgcgtg tgcgagaaaa tggccctcta cgacgtggtc 540
tccacccttc ctcaggctgt gatgggctcc tcatacggat tccagtactc tcctggacag 600
cgggtcgagt tcctggtgaa tgcctggaag tcaaagaaat gccctatggg cttcgcatat 660
gacactcgct gtttcgactc aacggtcacc gagaatgaca tccgtgttga ggagtcaatt 720

```
taccaatgtt gtgacttggc ccccgaagcc agacaggcca taaggtcgct tacagagcgg      780
ctttatatcg ggggtccccc gactaactca aaagggcaga gctgcggtta tcgccggtgc      840
cgtgcgagcg gcgtactgac gactagctgc ggtaataccc tcacatgtta cttgaaggcc      900
tctgcagcct gtcgagctgc aaagctccag gactgcacga tgcttgtgtg cggagacgac      960
ctcgtcgtta tctgtgaaag cgcgggaacc caagaggacg cggcgagcct acgagtcttc     1020
acggaggcta tgactaggta ctctgccccc cctggggacc cgccccaacc agaatacgac     1080
ttggagctaa taacatcatg ctcctccaac gtgtcggtcg cgcacgatgc atctggcaaa     1140
agggtatact acctcacccg tgaccccacc acccctcttg cgcgggctgc gtgggagaca     1200
gctagacaca ctccagttaa ctcctggcta ggcaacatca tcatgtatgc gcccactctg     1260
tgggcaagga tgattctgat gactcacttc ttctccatcc ttctagctca ggagcaactt     1320
gaaaaagccc tagattgtga gatctacggg gcccactact ccattgagcc acttgaccta     1380
cctcagatca ttcaacgact ccatggtctt agcgcgtttt cactccacag ttactctcca     1440
ggtgaaatca atagggtggc ttcatgcctc aggaagcttg gggtaccacc cttgcgagtc     1500
tggagacatc gggccagaag cgtccgcgct aagttactgt cccaggggggg gagggctgcc     1560
acttgtggca agtacctctt caactgggca gtaaggacca agcttaaact cactccaatt     1620
ccggctgcgc cccagttgga cttatccggc tggttcgttg ctggttacag cgggggagac     1680
atatatcaca gcctgtctcg tgcccgaccc cgctggttca tgtggtgcct actcctactt     1740
tctgtagggg taggcatcta cctgctcccc aaccgatga                            1779
```

<210> 8
<211> 592
<212> PRT
<213> Hepatitis C Virus

<400> 8

```
Met Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala
 1               5                  10                  15
Ala Glu Glu Ser Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu
            20                  25                  30
Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Ser Gln
        35                  40                  45
Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His
    50                  55                  60
Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Val Ser Thr Val Lys
65                  70                  75                  80
Ala Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro Pro His
                85                  90                  95
Ser Ala Arg Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu
            100                 105                 110
Ser Ser Lys Ala Val Asn His Ile His Ser Val Trp Lys Asp Leu Leu
        115                 120                 125
Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu
    130                 135                 140
Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu
145                 150                 155                 160
Ile Val Tyr Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu
                165                 170                 175
Tyr Asp Val Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser Ser Tyr
            180                 185                 190
Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Ala
        195                 200                 205
Trp Lys Ser Lys Lys Cys Pro Met Gly Phe Ala Tyr Asp Thr Arg Cys
    210                 215                 220
Phe Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile
225                 230                 235                 240
Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Arg Ser
                245                 250                 255
Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly
            260                 265                 270
Gln Ser Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr
        275                 280                 285
```

```
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys
    290                 295                 300
Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp
305                 310                 315                 320
Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser
                325                 330                 335
Leu Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly
            340                 345                 350
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser
        355                 360                 365
Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr
    370                 375                 380
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr
385                 390                 395                 400
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr
                405                 410                 415
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser
            420                 425                 430
Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Glu Ile
        435                 440                 445
Tyr Gly Ala His Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile
    450                 455                 460
Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro
465                 470                 475                 480
Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro
                485                 490                 495
Pro Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Lys Leu
            500                 505                 510
Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn
        515                 520                 525
Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser
    530                 535                 540
Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp
545                 550                 555                 560
Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Met Trp Cys
                565                 570                 575
Leu Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg
            580                 585                 590
```

<210> 9
<211> 592
<212> PRT
<213> Hepatitis C Virus

<400> 9

```
Met Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala
 1               5                  10                  15
Ala Glu Glu Ser Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu
            20                  25                  30
Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Ser Gln
        35                  40                  45
Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His
    50                  55                  60
Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Val Ser Thr Val Lys
65                  70                  75                  80
Ala Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro Pro His
                85                  90                  95
Ser Ala Arg Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu
            100                 105                 110
Ser Ser Lys Ala Val Asn His Ile His Ser Val Trp Lys Asp Leu Leu
        115                 120                 125
```

```
Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu
    130                 135                 140
Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu
145                 150                 155                 160
Ile Val Tyr Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu
                165                 170                 175
Tyr Asp Val Ile Ser Thr Leu Pro Gln Ala Val Met Gly Ser Ser Tyr
            180                 185                 190
Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Ala
        195                 200                 205
Trp Lys Ser Lys Lys Cys Pro Met Gly Phe Ala Tyr Asp Thr Arg Cys
    210                 215                 220
Phe Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile
225                 230                 235                 240
Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Arg Ser
                245                 250                 255
Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly
            260                 265                 270
Gln Ser Cys Gly His Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr
        275                 280                 285
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys
    290                 295                 300
Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp
305                 310                 315                 320
Leu Ile Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser
                325                 330                 335
Leu Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly
            340                 345                 350
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser
        355                 360                 365
Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr
    370                 375                 380
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr
385                 390                 395                 400
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr
                405                 410                 415
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser
            420                 425                 430
Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Glu Ile
        435                 440                 445
Tyr Gly Ala His Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile
    450                 455                 460
Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro
465                 470                 475                 480
Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro
                485                 490                 495
Pro Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Lys Leu
            500                 505                 510
Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn
        515                 520                 525
Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser
    530                 535                 540
Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp
545                 550                 555                 560
Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Cys Trp Phe Met Trp Cys
                565                 570                 575
Leu Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg
            580                 585                 590
```

<210> 10
<211> 592

<212> PRT
<213> Hepatitis C Virus

<400> 10

```
Met Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala
 1               5                  10                  15
Ala Glu Glu Ser Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu
            20                  25                  30
Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Ser Gln
        35                  40                  45
Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His
    50                  55                  60
Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Val Ser Thr Val Lys
65                  70                  75                  80
Ala Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro Pro His
                85                  90                  95
Ser Ala Arg Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu
            100                 105                 110
Ser Ser Lys Ala Val Asn His Ile His Ser Val Trp Lys Asp Leu Leu
        115                 120                 125
Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu
    130                 135                 140
Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu
145                 150                 155                 160
Ile Val Tyr Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu
                165                 170                 175
Tyr Asp Val Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser Ser Tyr
            180                 185                 190
Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Ala
        195                 200                 205
Trp Lys Pro Lys Lys Cys Pro Met Gly Phe Ala Tyr Asp Thr Arg Cys
    210                 215                 220
Phe Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile
225                 230                 235                 240
Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Arg Ser
                245                 250                 255
Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly
            260                 265                 270
Gln Ser Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr
        275                 280                 285
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys
    290                 295                 300
Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp
305                 310                 315                 320
Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser
                325                 330                 335
Leu Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly
            340                 345                 350
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser
        355                 360                 365
Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr
    370                 375                 380
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr
385                 390                 395                 400
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Thr Ile Met Tyr
                405                 410                 415
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser
            420                 425                 430
Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Glu Ile
        435                 440                 445
Tyr Gly Ala His Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile
```

```
   450                 455                 460
Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro
465                 470                 475                 480
Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro
                485                 490                 495
Pro Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Lys Leu
            500                 505                 510
Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn
        515                 520                 525
Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser
    530                 535                 540
Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp
545                 550                 555                 560
Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Met Trp Cys
                565                 570                 575
Leu Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg
            580                 585                 590
```

<210> 11
<211> 592
<212> PRT
<213> Hepatitis C Virus

<400> 11

```
Met Ser Met Ser Tyr Thr Arg Thr Gly Ala Leu Ile Thr Pro Cys Ala
 1               5                  10                  15
Ala Glu Glu Ser Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu
            20                  25                  30
Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Ser Gln
        35                  40                  45
Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His
    50                  55                  60
Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Val Ser Thr Val Lys
65                  70                  75                  80
Ala Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro Pro His
                85                  90                  95
Ser Ala Arg Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu
            100                 105                 110
Ser Ser Lys Ala Val Asn His Ile His Ser Val Trp Lys Asp Leu Leu
        115                 120                 125
Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu
    130                 135                 140
Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu
145                 150                 155                 160
Ile Val Tyr Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu
                165                 170                 175
Tyr Asp Val Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser Ser Tyr
            180                 185                 190
Gly Phe Gln His Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Ala
        195                 200                 205
Trp Lys Ser Lys Lys Cys Pro Met Gly Phe Ala Tyr Asp Thr Arg Cys
    210                 215                 220
Phe Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile
225                 230                 235                 240
Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Arg Ser
                245                 250                 255
Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly
            260                 265                 270
Gln Ser Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr
        275                 280                 285
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys
```

```
   290                 295                 300
Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp
305                 310                 315                 320
Phe Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser
                325                 330                 335
Leu Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly
            340                 345                 350
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser
        355                 360                 365
Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr
    370                 375                 380
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr
385                 390                 395                 400
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr
                405                 410                 415
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser
            420                 425                 430
Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Glu Ile
        435                 440                 445
Tyr Gly Ala His Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile
    450                 455                 460
Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro
465                 470                 475                 480
Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro
                485                 490                 495
Pro Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Lys Leu
            500                 505                 510
Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn
        515                 520                 525
Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser
    530                 535                 540
Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp
545                 550                 555                 560
Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Met Trp Cys
                565                 570                 575
Leu Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg
            580                 585                 590
```

<210> 12
<211> 592
<212> PRT
<213> Hepatitis C Virus

<400> 12

```
Val Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala
 1               5                  10                  15
Ala Glu Glu Ser Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu
            20                  25                  30
Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Ser Gln
        35                  40                  45
Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His
    50                  55                  60
Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Val Ser Thr Val Lys
65                  70                  75                  80
Ala Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro Pro His
                85                  90                  95
Ser Ala Arg Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu
            100                 105                 110
Ser Ser Lys Ala Val Asn His Ile His Ser Val Trp Lys Asp Leu Leu
        115                 120                 125
Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu
```

```
    130                 135                 140
Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu
145                 150                 155                 160
Ile Val Tyr Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu
                165                 170                 175
Tyr Asp Val Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser Ser Tyr
            180                 185                 190
Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Ala
        195                 200                 205
Trp Lys Ser Lys Lys Cys Pro Met Gly Phe Ala Tyr Asp Thr Arg Cys
    210                 215                 220
Phe Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile
225                 230                 235                 240
Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Arg Ser
                245                 250                 255
Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly
            260                 265                 270
Gln Ser Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr
        275                 280                 285
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys
    290                 295                 300
Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp
305                 310                 315                 320
Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser
                325                 330                 335
Leu Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly
            340                 345                 350
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser
        355                 360                 365
Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr
    370                 375                 380
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr
385                 390                 395                 400
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr
                405                 410                 415
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser
            420                 425                 430
Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Glu Ile
        435                 440                 445
Tyr Gly Ala His Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile
    450                 455                 460
Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro
465                 470                 475                 480
Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro
                485                 490                 495
Pro Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Lys Leu
            500                 505                 510
Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn
        515                 520                 525
Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Pro
    530                 535                 540
Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp
545                 550                 555                 560
Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Met Trp Cys
                565                 570                 575
Leu Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg
            580                 585                 590
```

<210> 13
<211> 592
<212> PRT

<213> Hepatitis C Virus

<400> 13

```
Met Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala
 1               5                  10                  15
Ala Glu Glu Ser Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu
            20                  25                  30
Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Ser Gln
        35                  40                  45
Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His
    50                  55                  60
Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Val Ser Thr Val Lys
65                  70                  75                  80
Ala Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro Pro His
                85                  90                  95
Ser Ala Arg Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu
            100                 105                 110
Ser Ser Lys Ala Val Asn His Ile His Ser Val Trp Lys Asp Leu Leu
        115                 120                 125
Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu
    130                 135                 140
Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu
145                 150                 155                 160
Ile Val Tyr Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu
                165                 170                 175
Tyr Asp Val Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser Ser Tyr
            180                 185                 190
Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Ala
        195                 200                 205
Trp Lys Ser Lys Lys Cys Pro Met Gly Phe Ala Tyr Asp Thr Arg Cys
    210                 215                 220
Phe Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile
225                 230                 235                 240
Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Arg Ser
                245                 250                 255
Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly
            260                 265                 270
Gln Ser Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr
        275                 280                 285
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys
    290                 295                 300
Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp
305                 310                 315                 320
Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser
                325                 330                 335
Leu Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly
            340                 345                 350
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser
        355                 360                 365
Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr
    370                 375                 380
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr
385                 390                 395                 400
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr
                405                 410                 415
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser
            420                 425                 430
Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Glu Ile
        435                 440                 445
Tyr Gly Ala His Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile
    450                 455                 460
```

```
Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro
465                 470                 475                 480
Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro
                485                 490                 495
Pro Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Lys Leu
            500                 505                 510
Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn
        515                 520                 525
Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Pro
    530                 535                 540
Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp
545                 550                 555                 560
Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Met Trp Cys
                565                 570                 575
Leu Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg
            580                 585                 590
```

<210> 14
<211> 20
<212> DNA
<213> Hepatitis C Virus

<400> 14
tcaatgtcct acacatggac 20

<210> 15
<211> 10
<212> DNA
<213> Hepatitis C Virus

<400> 15
ctacacatgg 10

<210> 16
<211> 41
<212> DNA
<213> Hepatitis C Virus

<400> 16
ctctgattac accatgcgct gcggaggaga gcaagctgcc c 41

<210> 17
<211> 28
<212> DNA
<213> Hepatitis C Virus

<400> 17
aatgcgctga gcaactcttt gctgcgcc 28

<210> 18
<211> 47
<212> DNA
<213> Hepatitis C Virus

<400> 18
ccataacatg gtctatgcca caacatcccg cagcgcaagc cagcggc 47

<210> 19
<211> 22
<212> DNA
<213> Hepatitis C Virus

<400> 19
gaagaaggta acttttgaca gg 22

<210> 20
<211> 23
<212> DNA
<213> Hepatitis C Virus

<400> 20
caagtcctgg atgaccacta ccg 23

<210> 21
<211> 14
<212> DNA
<213> Hepatitis C Virus

<400> 21
gacgtgctca agga 14

<210> 22
<211> 18
<212> DNA
<213> Hepatitis C Virus

<400> 22
atgaaggcga aggcgtcc 18

<210> 23
<211> 16
<212> DNA
<213> Hepatitis C Virus

<400> 23
ggaagaagcc tgtaag 16

<210> 24
<211> 24
<212> DNA
<213> Hepatitis C Virus

<400> 24
gaacctatcc agcaaggccg ttaa 24

<210> 25
<211> 19
<212> DNA
<213> Hepatitis C Virus

<400> 25
accagagaaa ggaggccgc 19

<210> 26
<211> 14
<212> DNA
<213> Hepatitis C Virus

<400> 26
acccagactt gggg 14

<210> 27
<211> 20

<212> DNA
<213> Hepatitis C Virus

<400> 27
tctccaccct tcctcaggct 20

<210> 28
<211> 19
<212> DNA
<213> Hepatitis C Virus

<400> 28
cgagttcctg gtgaatgcc 19

<210> 29
<211> 24
<212> DNA
<213> Hepatitis C Virus

<400> 29
tgccctatgg gcttcgcata tgac 24

<210> 30
<211> 25
<212> DNA
<213> Hepatitis C Virus

<400> 30
tttcgactca acggtcaccg agaat 25

<210> 31
<211> 15
<212> DNA
<213> Hepatitis C Virus

<400> 31
gttgaggagt caatt 15

<210> 32
<211> 20
<212> DNA
<213> Hepatitis C Virus

<400> 32
ttggcccccg aagccagaca 20

<210> 33
<211> 17
<212> DNA
<213> Hepatitis C Virus

<400> 33
aaggtcgctt acagagc 17

<210> 34
<211> 17
<212> DNA
<213> Hepatitis C Virus

<400> 34
atcgggggtc ccctgac 17

<210> 35
<211> 18
<212> DNA
<213> Hepatitis C Virus

<400> 35
taactcaaaa gggcagag 18

<210> 36
<211> 19
<212> DNA
<213> Hepatitis C Virus

<400> 36
atgttacttg aaggcctct 19

<210> 37
<211> 25
<212> DNA
<213> Hepatitis C Virus

<400> 37
gatgcttgtg tgcggagacg acctc 25

<210> 38
<211> 31
<212> DNA
<213> Hepatitis C Virus

<400> 38
ggtcgcgcac gatgcatctg gcaaaagggt a 31

<210> 39
<211> 19
<212> DNA
<213> Hepatitis C Virus

<400> 39
caccacccct cttgcgcgg 19

<210> 40
<211> 25
<212> DNA
<213> Hepatitis C Virus

<400> 40
ctccatcctt ctagctcagg agcaa 25

<210> 41
<211> 20
<212> DNA
<213> Hepatitis C Virus

<400> 41
agttactgtc ccaggggggg 20

<210> 42
<211> 17
<212> DNA
<213> Hepatitis C Virus

<400> 42
tccggctgcg tcccagt 17

<210> 43
<211> 26
<212> DNA
<213> Hepatitis C Virus

<400> 43
ccgaccccgc tggttcatgt ggtgcc 26

<210> 44
<211> 19
<212> DNA
<213> Hepatitis C Virus

<400> 44
ctacctgctc ccgaaccga 19

<210> 45
<211> 19
<212> DNA
<213> Hepatitis C Virus

<400> 45
ctacctgctc cccaaccga 19

<210> 46
<211> 80
<212> PRT
<213> cHepatitis C Virus

<400> 46

Ser	Met	Ser	Tyr	Thr	Trp	Thr	Gly	Ala	Leu	Ile	Thr	Pro	Cys	Ala	Ala
1				5					10					15	
Glu	Glu	Ser	Lys	Leu	Pro	Ile	Asn	Ala	Leu	Ser	Asn	Ser	Leu	Leu	Arg
			20					25				30			
His	His	Asn	Met	Val	Tyr	Ala	Thr	Thr	Ser	Arg	Ser	Ala	Ser	Gln	Arg
		35					40					45			
Gln	Lys	Lys	Val	Thr	Phe	Asp	Arg	Leu	Gln	Val	Leu	Asp	Asp	His	Tyr
	50					55					60				
Arg	Asp	Val	Leu	Lys	Glu	Met	Lys	Ala	Lys	Ala	Ser	Thr	Val	Lys	Ala
65					70					75					80

<210> 47
<211> 47
<212> PRT
<213> cHepatitis C Virus

<400> 47

Ser	Val	Glu	Glu	Ala	Cys	Lys	Leu	Thr	Pro	Pro	His	Ser	Ala	Arg	Ser
1				5					10					15	
Lys	Phe	Gly	Tyr	Gly	Ala	Lys	Asp	Val	Arg	Asn	Leu	Ser	Ser	Lys	Ala
			20					25				30			
Val	Asn	His	Ile	His	Ser	Val	Trp	Lys	Asp	Leu	Leu	Glu	Asp	Thr	
		35					40					45			

<210> 48
<211> 40
<212> PRT
<213> cHepatitis C Virus

<400> 48
Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys
1 5 10 15
Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile Val Tyr
20 25 30
Pro Asp Leu Gly Val Arg Val Cys
35 40

<210> 49
<211> 102
<212> PRT
<213> cHepatitis C Virus

<400> 49
Ser Ser Tyr Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu
1 5 10 15
Val Asn Ala Trp Lys Ser Lys Lys Cys Pro Met Gly Phe Ala Tyr Asp
20 25 30
Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu
35 40 45
Glu Ser Ile Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala
50 55 60
Ile Arg Ser Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn
65 70 75 80
Ser Lys Gly Gln Ser Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val
85 90 95
Leu Thr Thr Ser Cys Gly
100

<210> 50
<211> 56
<212> PRT
<213> cHepatitis C Virus

<400> 50
Cys Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Cys Glu Ser
1 5 10 15
Ala Gly Thr Gln Glu Asp Ala Ala Ser Leu Arg Val Phe Thr Glu Ala
20 25 30
Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Gln Pro Glu Tyr
35 40 45
Asp Leu Glu Leu Ile Thr Ser Cys
50 55

<210> 51
<211> 37
<212> PRT
<213> cHepatitis C Virus

<400> 51
Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr Leu Thr Arg
1 5 10 15
Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr Ala Arg His
20 25 30
Thr Pro Val Asn Ser
35

<210> 52
<211> 29
<212> PRT
<213> cHepatitis C Virus

<400> 52
Cys Leu Arg Lys Leu Gly Val Pro Pro Leu Arg Val Trp Arg His Arg
1 5 10 15
Ala Arg Ser Val Arg Ala Lys Leu Leu Ser Gln Gly Gly
20 25

<210> 53
<211> 25
<212> PRT
<213> Hepatitis C Virus

<400> 53
Phe Val Ala Gly Tyr Ser Gly Gly Asp Ile Tyr His Ser Leu Ser Arg
1 5 10 15
Ala Arg Pro Arg Trp Phe Met Trp Cys
20 25

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/07404

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :C12Q 1/70, 1/68; C12N 15/00; A61K 39/29
US CL : 424/189.1, 199.1; 435/5, 6, 320.1; 536/23.72
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/189.1, 199.1; 435/5, 6, 320.1; 536/23.72

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, CAPLUS, CA, WEST
search terms: Hepatitis C virus, NS5B, derivativ?, diagnostic, vaccine, immune response, antibodies, HCV.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,372,928 a (MIYAMURA et al) 13 December 1994, see the abstract, and figure 5.	1-15, 18, 19, 29, 31, 33, 35, 38-41, 90, 100
Y	US 5,371,017 A (HOUGHTON et al) 06 December 1994, column 3, lines 51-60.	1-15, 18, 19, 29, 31, 33, 35, 38-41, 90, 100
Y	BEHRENS et al. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. The EMBO Journal. 1996, Vol. 15, No. 1, pages 12-22, see the entire document.	1-15, 18, 19, 29, 31, 33, 35, 38-41, 90, 100

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
09 JULY 1999

Date of mailing of the international search report
22 JUL 1999

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
ALI R. SALIMI
Telephone No: (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/07404

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	LOHMANN et al. Biochemical Properties of Hepatitis C Virus NS5B RNA-Dependent RNA Polymerase and Identification of Amino Acid Sequence Motifs Essential for Enzymatic Activity. Journal of Virology. November 1997, Vol. 71, No. 1, pages 8416-8428, see the entire document.	1-15, 18, 19, 29, 31, 33, 35, 38-41, 90, 100
A	MILLER et al. Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. Proc. Natl. Acad. Sci. USA. March 1990, Vol. 87, pages 2057-2061.	1-15, 18, 19, 29, 31, 33, 35, 38-41, 90, 100

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/07404

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-15, 18, 19, 29, 31, 33, 35, 38-41, 90, 100

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/07404

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s)1-15, 18, 19, 29, 31, 33, 35, 38-41, 90, 100, drawn to nucleic acid encoding a hepatitis C virus (HCV) NS5B protein, variants , mutants and derivatives, and host cells.
Group II, claim(s) 16, 17, 20-26, 30, 32, 34, 36, 42, drawn to hepatitis C virus proteins.
Group III(a)-(n), claim(s) 27, 28, 44, drawn to hepatitis V virus proteins, alterations of the proteins and various mutants.
Group IV(a)-(k), claim(s)37, 43, 81, 91, 92, 93, 94-97, drawn to mutants of various nucleic acids of hepatitis C virus plus variants and multiple altered derivatives.
Group V, claim(s) 45-50, drawn to method of assaying a test compound for modulating activity against HCV.
Group VI, claim(s) 51-55, drawn to method of assaying a compound for antagonistic activity.
Group VII, claim(s)56-60, drawn to method of assaying a compound for agonistic activity against HCV.
Group VIII, claim(s) 61-65, drawn to method of assaying a test compound for interaction with HCV protein.
Group IX, claim(s) 66-67, drawn to method of interaction with HCV nucleic acid.
Group X, claim(s)68-71, drawn to method of detecting HCV in a biological sample comprising nucleic acid amplification.
Group XI, claim(s) 72-75, drawn to method for detecting interactions between viral polypeptide and antibodies.
Group XII(a)-(h), claim(s) 76-79, drawn to antibodies.
Group XIII, claim(s)80, drawn to antibody with specific affinity to polypeptide comprising in Table 2.
Group XIV, claim(s) 82, 87, drawn to method of introducing a nucleic acid into a host.
Group XV, claim(s) 83-86, drawn to viral antigens.
Group XVI, claim(s)88, drawn to method of inducing an immune response.
Group XVII, claim(s) 89, drawn to a method of amplifying HCV nucleic acids.
Group XVIII, claim(s) 98, 99, drawn to method of assessing the functionality of HCV NS5B protein.
Group XIX, claim(s)101, drawn to a method of assaying a test compound for antagonist against HCV in a cell line.
Group XX, claim(s) 102, drawn to a method of assaying a test compound for agonist activity HCV in a cell line.
Group XXI, claim(s) 103-110, drawn to a method of preparing a HCV NS5B protein.
Group XXII, claim(s)111, drawn to a method of isolating an NS5B protein.
Group XXIII, claim(s) 112, 113, 119, 120, drawn to a method of generating an infectious viral vector.
Group XXIV, claim(s) 114, 115, drawn to a method of generating an infectious viral vector comprising substituting HCV NS5B homologous gene, and modified genes as set forth in Table 6.

Group XXV, claim(s)116, drawn to a host animal.
Group XXVI, claim(s) 117, drawn to a method of isolating antibodies by screening antibody library.
Group XXVII, claim(s) 118, drawn to a method of propagating HCV in a in vitro system utilizing SEQ ID NO: 1-7.
Group XXVIII, claim(s)121-124, drawn to method of propagating HCV in an in vitro system utilizing SEQ ID NO: 2, 8-13.
Group XXIX, claim(s) 125-128, drawn to a method of propagating HCV in a living host.
Group XXX, claim(s) 129, 130, drawn to a kit for detecting the HCV nucleic acids.
Group XXXI, claim(s)131-132, drawn to a kit for detecting HCV utilizing antibodies.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows: Nucleic acids molecule as set forth in Table 6 (species 1-11)as it applies to claims of Group IV. Polypeptide of Table 4, 5 (total species 1-14)as they apply to claims of Group III. Antibodies for polypeptide of Table 2 (1-8). For explanation see below.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/07404

The inventions listed as Groups I-XXXI do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The invention of Group I does not make a contribution over the prior art as evidence by Lohmann et al (Journal of Virology, Nov. 1997, Vol. 71, No. 11, p. 8416-8428), also as evidence by Miyamura et al (US Patent No. 5,372,928, Dec. 1994), and Behrens et al (The EMBO Journal, 1996, Vol. 15, No. 1, p. 12-22). Since the nucleic acid molecule of hepatitis C virus NS5B is already known in the art thus the invention lacks unity of invention as defined by PCT Rule 13.2. The cited references prove that the technical feature of Group I does not make a contribution over the prior art. The claims are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a single inventive concept, accordingly, the unity of invention is lacking among all groups.

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons: Each mutations listed confers different structure and presumably different effect on antigenicity and virulence of the virus. Since the effects of mutations are unpredictable, a mutation of one location does not teach or suggests mutation at a different location. Therefore, each location is distinct species.